

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 January 2002 (03.01.2002)

PCT

(10) International Publication Number  
WO 02/00722 A2

(51) International Patent Classification: C07K 14/705 (81) Designated States (optional): AE, AG, AL, AM, AT, AU, A2, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EG, ES, FI, FR, GB, GR, GU, GM, HR, HU, ID, IL, IN, JP, KE, KG, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NG, NZ, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(21) International Application Number: PCT/US01/20640

(22) International Filing Date: 26 June 2001 (26.06.2001)

(23) Filing Language: English

(24) Publication Language: English

(30) Priority Date: 26 June 2000 (26.06.2000) US 60214176

(71) Applicant (for all designated States except US): MIL-LENNIUM PHARMACEUTICALS, INC. (US05): 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors and (75) Inventor/Applicants (for US only): CURTIS, Rory, A., J. (CH/US); 31 Constitution Drive, Southborough, MA 01772 (US); SILOS-SANTIAGO, Inmaculada (ES/US); 18 Hillland Street, Cambridge, MA 02138 (US).

(74) Agents: NANNARAGOURAS, Amy, E.; Jahnke & Cocke, Ltd., 28 State Street, Boston, MA 02109 et al. (US).

Published: — without international search report and to be republished upon receipt of that report  
For prior art purposes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/00722 A2

(54) Title: 48000 AND 52920, NOVEL HUMAN CALCIUM CHANNELS AND USES THEREOF

(57) Abstract: The invention provides isolated nucleic acid molecules, designated TLCC-4 and TLCC-5 nucleic acid molecules, which encode novel TRP-like calcium channel molecules. The invention also provides anti-sense nucleic acid molecules, recombinant expression vectors containing TLCC-4 or TLCC-5 nucleic acid molecules, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a TLCC-4 or TLCC-5 gene has been introduced or disrupted. The invention still further provides isolated TLCC-4 or TLCC-5 polypeptides, fusion polypeptides, antigenic peptides and anti-TLCC-4 or anti-TLCC-5 antibodies. Diagnostic methods utilizing compositions of the invention are also provided.

WO 02/00722 PCT/US01/20640

48000 AND 52920, NOVEL HUMAN CALCIUM CHANNELS AND USES THEREOF

Related Applications

This application claims priority to U.S. Provisional Patent Application No. 60/214,176 filed on June 26, 2000, incorporated herein in its entirety by reference.

Background of the Invention

Calcium signaling has been implicated in the regulation of a variety of cellular responses, such as growth and differentiation. There are two general methods by which intracellular concentrations of calcium ions may be increased: calcium ions may be freed from intracellular stores, transported by specific membrane channels in the storage organelle, or calcium ions may be brought into the cell from the extracellular milieu through the use of specific channels in the cellular membrane. In the situation in which the intracellular stores of calcium have been depleted, a specific type of calcium channel, termed a 'capacitative calcium channel' or a 'store-operated calcium channel' (SOC), is activated in the plasma membrane to import calcium ions from the extracellular environment to the cytosol (see Putney and McKay (1999) *BioEssays* 21:38-46). Calcium may also enter the cell via receptor-stimulated cation channels (see Hofmann *et al.* (2000) *J. Mol. Med.* 78:14-25).

There is no single electrophysiological profile characteristic of the calcium channel family; rather, a wide array of single channel conductances, cation selectivity, and current properties have been observed for different channels. Further, in several instances it has been demonstrated that homo- or hetero-polymerization of the channel molecule may occur, further changing the channel properties from those of the single molecule. In general, though, these channels function similarly, in that they are calcium ion-permeable cation channels which become activated after agonist binding to a G protein-coupled receptor.

Members of the capacitative calcium channel family include the calcium release-activated calcium current (CRAC) (Hoth and Penner (1992) *Nature* 355: 353-355), calcium release-activated non-selective cation current (CRANC) (Krause *et al.* (1996) *J. Biol. Chem.* 271: 32523-32528), and the transient receptor potential (TRP) proteins TRP1, TRP2, TRP4, and TRP5. Depletion of intracellular calcium stores activate these channels by a mechanism which is yet undefined, but which has been demonstrated to involve a diffusible factor using studies in which calcium stores were artificially depleted (e.g., by the introduction of chelators into the cell, by activating phospholipase C<sub>γ</sub> or by inhibiting those enzymes responsible for pumping calcium ions into the stores or those enzymes responsible for maintaining resting intracellular calcium ion

concentrations) (Putney, J.W., (1986) *Cell Calcium* 7:1-12; Putney, J.W. (1990) *Cell Calcium* 11:611-624).

Recently, it has been elucidated that three TRP family members, TRP3, TRP6, and a mouse homologue, TRP7, form a sub-family of receptors that are activated in a calcium store-depletion independent manner. TRP3 and TRP6 are activated by diacylglycerols in a membrane delimited manner (Hofmann *et al.* (1999) *Nature* 397:259-263). Similarly, murine TRP7 is activated via diacylglycerol stimulation by G<sub>q</sub> protein coupled receptors (Okada *et al.* (1999) *J. Biol. Chem.* 274:27359-27370).

The TRP channel family is one of the best characterized calcium channel protein families. These channels include transient receptor potential protein and homologues thereof (to date, seven TRP homologues and splice variants have been identified in a variety of organisms), the vanilloid receptor subtype 1 (also known as the capsaicin receptor); the stretch-inhibitable non-selective cation channel (SIC); the olfactory, mechanosensitive channel; the insulin-like growth factor I-regulated calcium channel; the vitamin D-responsive apical, epithelial calcium channel (ECaC); melastatin; and the polycystic kidney disease protein family (see, *e.g.*, Montell and Rubin (1989) *Neuron* 2:1313-1323; Caterina *et al.* (1997) *Nature* 389: 816-824; Suzuki *et al.* (1999) *J. Biol. Chem.* 274: 6330-6335; Kiselyov *et al.* (1998) *Nature* 396: 478-482; Hoenderop *et al.* (1999) *J. Biol. Chem.* 274: 8375-8378; and Chen *et al.* (1999) *Nature* 401(6751): 383-6).

Each of these molecules is 700 or more amino acids in length, and shares certain conserved structural features. Predominant among these structural features are six transmembrane domains, with an additional hydrophobic loop present between the fifth and sixth transmembrane domains. It is believed that this loop is integral to the activity of the pore of the channel formed upon membrane insertion (Hardie and Minke (1993) *Trends Neurosci* 16: 371-376). TRP channel proteins also include one or more ankyrin domains. Although found in disparate tissues and organisms, members of the TRP channel protein family all serve to transduce signals by means of calcium entry into cells, particularly pain signals (see, *e.g.*, McClesky and Gold (1999) *Annu. Rev. Physiol.* 61: 835-856), light signals (Hardie and Minke, *supra*), or olfactory signals (Colbert *et al.* (1997) *J. Neurosci* 17(21): 8259-8269). Thus, this family of molecules may play important roles in sensory signal transduction in general.

Vanilloid receptors (VRs) are cation channels that are structurally related to members of the TRP family of ion channels. VRs share several physical characteristics including an N-terminal cytoplasmic domain which contains three ankyrin repeats, six transmembrane domains, a pore-loop region located between transmembrane domains 5 and 6, and several kinase consensus sequences. These receptors have been proposed to mediate the entry of extracellular calcium into cells in response to the depletion of intracellular calcium stores. VRs are expressed in nociceptive neurons, as well as other

cells types, and are activated by a variety of stimuli including noxious heat and protons. A well-known agonist of VR1 is capsaicin, which induces pain behavior in humans and rodents. VR1 knockout mice have been shown to be impaired in their detection of painful heat, to exhibit no vanilloid-evoked pain behavior, and to show little thermal hypersensitivity after inflammation (Szallasi and Blumberg (1999) *Pharmacol. Rev.* 51:159-211).

### Summary of the Invention

The present invention is based, at least in part, on the discovery of novel calcium channel family members, referred to herein as "TRP-like calcium channel-4" and "TRP-like calcium channel-5" or "TLCC-4" and "TLCC-5" nucleic acid and polypeptide molecules. The TLCC-4 and TLCC-5 nucleic acid and polypeptide molecules of the present invention are useful as modulating agents in regulating a variety of cellular processes, *e.g.*, membrane excitability, neurite outgrowth and synaptogenesis, signal transduction, cell proliferation, growth, differentiation, and migration, and nociception. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding TLCC-4 and TLCC-5 polypeptides or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of TLCC-4-encoding and TLCC-5-encoding nucleic acids.

In one embodiment, the invention features an isolated nucleic acid molecule that includes the nucleotide sequence set forth in SEQ ID NO:1, 3, 4, or 6. In another embodiment, the invention features an isolated nucleic acid molecule that encodes a polypeptide including the amino acid sequence set forth in SEQ ID NO:2 or 5. In another embodiment, the invention features isolated nucleic acid molecules that include the nucleotide sequences contained in the plasmids deposited with ATCC® as

### Accession Numbers

In still other embodiments, the invention features isolated nucleic acid molecules including nucleotide sequences that are substantially identical (*e.g.*, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 98% identical) to the nucleotide sequence set forth as SEQ ID NO:1, 3, 4, or 6. The invention further features isolated nucleic acid molecules including at least 50 contiguous nucleotides of the nucleotide sequence set forth as SEQ ID NO:1, 3, 4, or 6. In another embodiment, the invention features isolated nucleic acid molecules which encode a polypeptide including an amino acid sequence that is substantially identical (*e.g.*, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 98% identical) to the amino acid sequence set forth as SEQ ID NO:2 or 5. The present invention also features nucleic acid molecules which encode allelic variants of the polypeptide having the amino acid sequence set forth as SEQ ID NO:2 or 5. In addition to isolated nucleic acid molecules encoding full-length polypeptides, the present

invention also features nucleic acid molecules which encode fragments, for example biologically active or antigenic fragments, of the full-length polypeptides of the present invention (e.g., fragments including at least 10, 15, 20, 25, 30, 35, 40, 45 or 50 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2 or 5). In still other embodiments, the invention features nucleic acid molecules that are complementary to, antisense to, or hybridize under stringent conditions to the isolated nucleic acid molecules described herein.

In another aspect, the invention provides vectors including the isolated nucleic acid molecules described herein (e.g., TLCC-4-encoding or TLCC-5-encoding nucleic acid molecules). Such vectors can optionally include nucleotide sequences encoding heterologous polypeptides. Also featured are host cells including such vectors (e.g., host cells including vectors suitable for producing TLCC-4 or TLCC-5 nucleic acid molecules and polypeptides).

In another aspect, the invention features isolated TLCC-4 or TLCC-5 polypeptides and/or biologically active or antigenic fragments thereof. Exemplary embodiments feature a polypeptide including the amino acid sequence set forth as SEQ ID NO:2 or 5, a polypeptide including an amino acid sequence at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 98% identical to the amino acid sequence set forth as SEQ ID NO:2 or 5, a polypeptide encoded by a nucleic acid molecule including a nucleotide sequence at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 98% identical to the nucleotide sequence set forth as SEQ ID NO:1, 3, 4, or 6. Also featured are fragments of the full-length polypeptides described herein (e.g., fragments including at least 10 contiguous amino acid residues of the sequence set forth as SEQ ID NO:2 or 5) as well as allelic variants of the polypeptide having the amino acid sequence set forth as SEQ ID NO:2 or 5.

The TLCC-4 or TLCC-5 polypeptides and/or biologically active or antigenic fragments thereof, are useful, for example, as reagents or targets in assays applicable to treatment and/or diagnosis of TLCC-4 or TLCC-5 mediated or related disorders. In one embodiment, a TLCC-4 or TLCC-5 polypeptide, or fragment thereof, has a TLCC-4 or TLCC-5 activity. In another embodiment, a TLCC-4 or TLCC-5 polypeptide, or fragment thereof, has one or more of the following domains: an ankyrin repeat domain, a transmembrane domain, a pore domain, a transient receptor domain, a kinase consensus sequence, and/or an ion transport protein domain, and optionally, has a TLCC-4 or TLCC-5 activity. In a related aspect, the invention features antibodies (e.g., antibodies which specifically bind to any one of the polypeptides described herein) as well as fusion polypeptides including all or a fragment of a polypeptide described herein.

4

The present invention further features methods for detecting TLCC-4 or TLCC-5 polypeptides and/or TLCC-4 or TLCC-5 nucleic acid molecules, such methods featuring, for example, a probe, primer or antibody described herein. Also featured are kits, e.g., kits for the detection of TLCC-4 or TLCC-5 polypeptides and/or TLCC-4 or TLCC-5 nucleic acid molecules. In a related aspect, the invention features methods for identifying compounds which bind to and/or modulate the activity of a TLCC-4 or TLCC-5 polypeptide or TLCC-4 or TLCC-5 nucleic acid molecule described herein. Further featured are methods for modulating a TLCC-4 or TLCC-5 activity.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

#### Brief Description of the Drawings

*Figure 1* depicts the cDNA sequence and predicted amino acid sequence of human TLCC-4. The nucleotide sequence corresponds to nucleic acids 1 to 4586 of SEQ ID NO:1. The amino acid sequence corresponds to amino acids 1 to 742 of SEQ ID NO:2. The coding region without the 5' and 3' untranslated regions of the human TLCC-4 gene is shown in SEQ ID NO:3.

*Figure 2* depicts the cDNA sequence and predicted amino acid sequence of human TLCC-5. The nucleotide sequence corresponds to nucleic acids 1 to 3042 of SEQ ID NO:4. The amino acid sequence corresponds to amino acids 1 to 1013 of SEQ ID NO:5. The coding region of the human TLCC-5 gene is shown in SEQ ID NO:6.

*Figure 3* depicts TLCC-4 mRNA expression in normal human tissues, as determined using RT-PCR.

*Figure 4* depicts a structural, hydrophobicity, and antigenicity analysis of the human TLCC-4 polypeptide (SEQ ID NO:2).

*Figure 5* depicts the results of a search which was performed against the HMM database in PFAM and which resulted in the identification of three "ankyrin repeat domains" and an "ion transport protein domain" in the human TLCC-4 polypeptide (SEQ ID NO:2).

*Figure 6* depicts the results of a search which was performed against the MEMSAT database and which resulted in the identification of six "transmembrane domains" in the human TLCC-4 polypeptide (SEQ ID NO:2).

*Figure 7* depicts a structural, hydrophobicity, and antigenicity analysis of the human TLCC-5 polypeptide (SEQ ID NO:5).

*Figure 8* depicts the results of a search which was performed against the HMM database which resulted in the identification of two "transient receptor domains" in the amino acid sequence of the human TLCC-5 polypeptide (SEQ ID NO:5).

5

Figure 9 depicts the results of a search which was performed against the MEMSAT database and which resulted in the identification of two "transmembrane domains" in the human TLCC-5 polypeptide (SEQ ID NO:5).

Figure 10 depicts an alignment of the human TLCC-5 amino acid sequence (SEQ ID NO:5) with the amino acid sequence of human transient receptor 7 (hTRP7), using the CLUSTAL W (1.74) alignment program.

Figure 11 depicts an alignment of the human TLCC-5 amino acid sequence (SEQ ID NO:5) with the amino acid sequence of human melastatin, using the CLUSTAL W (1.74) alignment program.

Figure 12 is a graph depicting the expression of human TLCC-4 cDNA (SEQ ID NO:3) in various human tissues as determined by Taqman analysis.

Figure 13 is a graph depicting the expression of human TLCC-4 cDNA (SEQ ID NO:3) in various diseased and normal state human tissues as determined by Taqman analysis.

Figure 14 is a graph depicting the expression of human TLCC-4 cDNA (SEQ ID NO:3) in normal monkey and human tissues as determined by Taqman analysis.

Figure 15 is a graph depicting the expression of human TLCC-4 cDNA (SEQ ID NO:3) in normal monkey and human tissues as determined by Taqman analysis.

Figure 16 is a graph depicting the expression of human TLCC-4 cDNA (SEQ ID NO:3) in various tissues involved in metabolism as determined by Taqman analysis.

#### Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "TRP-like calcium channel-4" and "TRP-like calcium channel-5" or "TLCC-4" and "TLCC-5" nucleic acid and polypeptide molecules, which are novel members of the ion channel, *e.g.*, calcium channel and/or vanilloid receptor, family. These novel molecules are capable of, for example, modulating an ion-channel mediated activity (*e.g.*, a calcium channel- and/or vanilloid receptor-mediated activity) in a cell, *e.g.*, a neuronal, skin, muscle (*e.g.*, cardiac muscle), or liver cell.

As used herein, an "ion channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting signals in an electrically excitable cell, *e.g.*, a neuronal or muscle cell. Ion channels include calcium channels, potassium channels, and sodium channels. As used herein, a "calcium channel" includes a protein or polypeptide which is involved in receiving, conducting, and transmitting calcium ion-based signals in an electrically excitable cell. Calcium channels are calcium ion selective, and can determine membrane excitability (the ability of, for example, a neuronal cell to respond to a stimulus and to convert it into a sensory impulse). Calcium channels can also influence the resting potential of membranes, wave forms and

frequencies of action potentials, and thresholds of excitation. Calcium channels are typically expressed in electrically excitable cells, *e.g.*, neuronal cells, and may form heteromultimeric structures (*e.g.*, composed of more than one type of subunit). Calcium channels may also be found in non-excitable cells (*e.g.*, adipose cells or liver cells), where they may play a role in, *e.g.*, signal transduction. Examples of calcium channels include the low-voltage-gated channels and the high-voltage-gated channels. Calcium channels are described in, for example, Davila *et al.* (1999) *Annals New York Academy of Sciences* 868:102-17 and McEnery, M.W. *et al.* (1998) *J. Bioenergetics and Biomembranes* 30(4): 409-418, the contents of which are incorporated herein by reference. As the TLCC-4 and TLCC-5 molecules of the present invention are calcium channels modulating ion channel mediated activities (*e.g.*, calcium channel- and/or vanilloid receptor-mediated activities), they may be useful for developing novel diagnostic and therapeutic agents for ion channel associated disorders (*e.g.*, calcium channel and/or vanilloid receptor associated disorders).

As used herein, an "ion channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of an ion channel mediated activity. For example, a "calcium channel associated disorder" includes a disorder, disease or condition which is characterized by a misregulation of a calcium channel mediated activity. Ion channel associated disorders, *e.g.*, calcium channel associated disorders, include CNS disorders, such as cognitive and neurodegenerative disorders, examples of which include, but are not limited to, Alzheimer's disease, dementias related to Alzheimer's disease (such as Pick's disease), Parkinson's and other Lewy diffuse body diseases, senile dementia, Huntington's disease, Gilles de la Tourette's syndrome, multiple sclerosis, amyotrophic lateral sclerosis, progressive supranuclear palsy, epilepsy, Creutzfeldt-Jakob disease, or AIDS related dementia; autonomic function disorders such as hypertension and sleep disorders, and neuropsychiatric disorders, such as depression, schizophrenia, schizoaffective disorder, Korsakoff's psychosis, mania, anxiety disorders, or phobic disorders; learning or memory disorders, *e.g.*, amnesia or age-related memory loss, attention deficit disorder, psychoactive substance use disorders, anxiety, phobias, panic disorder, as well as bipolar affective disorder, *e.g.*, severe bipolar affective (mood) disorder (BP-1), and bipolar affective neurological disorders, *e.g.*, migraine and obesity. Further CNS-related disorders include, for example, those listed in the American Psychiatric Association's Diagnostic and Statistical manual of Mental Disorders (DSM), the most current version of which is incorporated herein by reference in its entirety.

Ion channel associated disorders, *e.g.*, vanilloid receptor associated disorders also include pain disorders. As used herein, the term "pain disorders" includes those disorders, diseases or conditions that affect pain signaling mechanisms. Examples of

pain disorders include postherpetic neuralgia, diabetic neuropathy, postmastectomy pain syndrome, stump pain, reflex sympathetic dystrophy, trigeminal neuralgia, neuropathic pain, orofacial neuropathic pain, osteoarthritis, rheumatoid arthritis, fibromyalgia syndrome, tension myalgia, Guillain-Barre syndrome, Meralgia paraesthetica, burning mouth syndrome, fibrocitis, myofascial pain syndrome, idiopathic pain disorder, temporomandibular joint syndrome, atypical odonalgia, loin pain, haematoma syndrome, non-cardiac chest pain, low back pain, chronic nonspecific pain, psychogenic pain, musculoskeletal pain disorder, chronic pelvic pain, nonorganic chronic headache, tension-type headache, cluster headache, migraine, complex regional pain syndrome, vaginismus, nerve trunk pain, somatoform pain disorder, cyclical mastalgia, chronic fatigue syndrome, multiple somatization syndrome, chronic pain disorder, somatization disorder, Syndrome X, facial pain, idiopathic pain disorder, posttraumatic rheumatic pain modulation disorder (fibrositis syndrome), and Tanager disease.

As used herein, the term "pain signaling mechanisms" include the cellular mechanisms involved in the development and regulation of pain, *e.g.*, pain elicited by noxious chemical, mechanical, or thermal stimuli, in a subject, *e.g.*, a mammal such as a human. In mammals, the initial detection of noxious chemical, mechanical, or thermal stimuli, a process referred to as "nociception", occurs predominantly at the peripheral terminals of specialized, small diameter sensory neurons. These sensory neurons transmit the information to the central nervous system, evoking a perception of pain or discomfort and initiating appropriate protective reflexes. The TLCC-4 or TLCC-5 molecules of the present invention may be present on these sensory neurons and, thus, may be involved in detecting these noxious chemical, mechanical, or thermal stimuli and transducing this information into membrane depolarization events. Thus, the TLCC-4 or TLCC-5 molecules by participating in pain signaling mechanisms, may modulate pain elicitation and act as targets for developing novel diagnostic targets and therapeutic agents to control pain.

Ion channel associated disorders, *e.g.*, calcium channel and/or vanilloid receptor disorders, also include cellular proliferation, growth, differentiation, or migration disorders. Cellular proliferation, growth, differentiation, or migration disorders include those disorders that affect cell proliferation, growth, differentiation, or migration processes. As used herein, a "cellular proliferation, growth, differentiation, or migration process" is a process by which a cell increases in number, size or content, by which a cell develops a specialized set of characteristics which differ from that of other cells, or by which a cell moves closer to or further from a particular location or stimulus. The TLCC-4 or TLCC-5 molecules of the present invention are involved in signal transduction mechanisms, which are known to be involved in cellular growth, differentiation, and migration processes. Thus, the TLCC-4 or TLCC-5 molecules may

modulate cellular growth, differentiation, or migration, and may play a role in disorders characterized by aberrantly regulated growth, differentiation, or migration. Such disorders include cancer, *e.g.*, carcinoma, sarcoma, or leukemia, tumor angiogenesis and metastasis, skeletal dysplasia, neuronal deficiencies resulting from impaired neural induction and patterning; hepatic disorders, cardiovascular disorders, and hematopoietic and/or myeloproliferative disorders.

As used herein, an "ion channel mediated activity" includes an activity which involves an ion channel, *e.g.*, an ion channel and/or a vanilloid receptor, in a neuronal cell, a muscular cell, a skin cell or a liver cell, associated with receiving, conducting, and transmitting signals. Ion channel mediated activities (*e.g.*, calcium channel and/or vanilloid receptor mediated activities) include release of neurotransmitters or second messenger molecules (*e.g.*, dopamine or norepinephrine), from cells, *e.g.*, neuronal cells; modulation of resting potential of membranes, wave forms and frequencies of action potentials, and thresholds of excitation; participation in signal transduction pathways, and modulation of processes such as integration of sub-threshold synaptic responses and the conductance of back-propagating action potentials in, for example, neuronal cells (*e.g.*, changes in those action potentials resulting in a morphological or differentiative response in the cell).

The term "family" when referring to the polypeptide and nucleic acid molecules of the invention is intended to mean two or more polypeptides or nucleic acid molecules having a common structural domain or motif and having sufficient amino acid or nucleotide sequence homology as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin, *e.g.*, monkey proteins. Members of a family may also have common functional characteristics.

For example, the family of TLCC-4 and TLCC-5 polypeptides comprise at least one "transmembrane domain." As used herein, the term "transmembrane domain" includes an amino acid sequence of about 20-45 amino acid residues in length which spans the plasma membrane. More preferably, a transmembrane domain includes about at least 20, 25, 30, 35, 40, or 45 amino acid residues and spans the plasma membrane. Transmembrane domains are rich in hydrophobic residues, and typically have an alpha-helical structure. In a preferred embodiment, at least 50%, 60%, 70%, 80%, 90%, 95% or more of the amino acids of a transmembrane domain are hydrophobic, *e.g.*, leucines, isoleucines, alanines, valines, phenylalanines, prolines or methionines. Transmembrane domains are described in, for example, Zagotta W.N. *et al.* (1996) *Annual Rev. Neurosci.* 19: 235-263, the contents of which are incorporated herein by reference.

Amino acid residues 440-461, 488-508, 520-540, 547-565, 590-609, and 652-676 of the human TLCC-4 polypeptide comprise transmembrane domains (see, e.g., Figures 4 and 6). Amino acid residues 786-803 and 826-848 of the human TLCC-5 polypeptide comprise transmembrane domains (see Figures 7 and 9). Accordingly, TLCC-4 and/or TLCC-5 polypeptides having at least 50-60% homology, preferably about 60-70%, more preferably about 70-80%, or about 80-90% homology with a transmembrane domain of human TLCC-4 and/or TLCC-5 are within the scope of the invention.

In another embodiment, a TLCC-4 molecule of the present invention is identified based on the presence of at least one "pore domain" between the fifth and sixth transmembrane domains. As used herein, the term "pore domain" includes an overall hydrophobic amino acid sequence which is located between two transmembrane domains of a calcium channel protein, preferably transmembrane domains 5 and 6, and which is believed to be a major determinant of ion selectivity and channel activity in calcium channels. Pore domains are described in, for example Vannier *et al.* (1998) *J. Biol. Chem.* 273: 8675-8679 and Phillips, A. M. *et al.* (1992) *Neuron* 8, 631-642, the contents of which are incorporated herein by reference. TLCC-4 molecules having at least one pore domain are within the scope of the invention. Amino acid residues 620-640 of the human TLCC-4 sequence (SEQ ID NO:2) comprise a pore domain (Figure 4).

In another embodiment, a TLCC-4 molecule of the present invention is identified based on the presence of at least one "ankyrin repeat domain." As used herein, the term "ankyrin repeat domain" includes an amino acid sequence of about 10-110 amino acid residues which serves as an ankyrin repeat. Preferably, an ankyrin repeat domain includes at least about 30 amino acid residues. To identify the presence of an ion transport domain in a TLCC-4 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the HMM database). The ankyrin repeat domain (HMM) has been assigned the PFAM Accession PF000023 (at the PFAM website). A search was performed against the HMM database resulting in the identification of ankyrin repeat domains in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 167-202, 214-246, 261-294, and 340-372 of SEQ ID NO:2. The results of the search are set forth in Figure 5.

In another embodiment, a TLCC-4 molecule of the present invention is identified based on the presence of at least one "ion transport protein domain." As used herein, the term "ion transport protein domain" includes a protein domain having an amino acid sequence of about 100-200 amino acid residues which serves to transport ions. Preferably, an ion transport protein domain includes at least about 160 amino acid residues. To identify the presence of an ion transport protein domain in a TLCC-4 protein, and make the determination that a protein of interest has a particular profile, the

amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the HMM database). The ion transport domain (HMM) has been assigned the PFAM Accession PF00520 (at the world wide web address: genome.wustl.edu/Pfam/html). A search was performed against the HMM database

resulting in the identification of an ion transport protein domain in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 510-677 of SEQ ID NO:2. The results of the search are set forth in Figure 5.

In another embodiment, a TLCC-5 molecule of the present invention is identified based on the presence of at least one "transient receptor domain." As used herein, the term "transient receptor domain" includes a protein domain having an amino acid sequence of about 100-200 amino acid residues which is found in transient receptor potential (Trp) proteins and related ion channel proteins. Preferably, a transient receptor domain includes at least about 56-58 amino acid residues. To identify the presence of a transient receptor domain in a TLCC-5 protein, and make the determination that a protein of interest has a particular profile, the amino acid sequence of the protein may be searched against a database of known protein domains (e.g., the HMM database). The transient receptor domain (HMM) has been assigned the PFAM Accession PF02164 (at the world wide web address: genome.wustl.edu/Pfam/html). A search was performed against the HMM database resulting in the identification of transient receptor domains in the amino acid sequence of human TLCC-5 (SEQ ID NO:5) at about residues 720-778 and 820-876 of SEQ ID NO:5. The results of the search are set forth in Figure 8.

A description of the Pfam database can be found in Sonhammer *et al.* (1997) *Proteins* 28:405-420 and a detailed description of HMMs can be found, for example, in Gribskov *et al.* (1990) *Meth. Enzymol.* 183:146-159; Gribskov *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh *et al.* (1994) *J. Mol. Biol.* 235:1501-1531; and Stultz *et al.* (1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference.

In a preferred embodiment, the TLCC-4 or TLCC-5 molecules of the invention include at least one transmembrane domain, at least one ankyrin repeat domain, at least one pore domain, at least one transient receptor domain, and/or at least one ion transport protein domain.

In a preferred embodiment, a TLCC-4 or TLCC-5 polypeptide includes at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous or identical to the amino acid sequence of SEQ ID NO:2 or 5, or the amino acid sequences encoded by the DNA inserts of the plasmids deposited with ATCC as

Accession Numbers \_\_\_\_\_. In yet another preferred embodiment, a TLCC-4 or TLCC-5 polypeptide includes at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6. In another preferred embodiment, a TLCC-4 or TLCC-5 polypeptide includes at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a pore domain, and/or a transient receptor domain, and/or an ion transport protein domain, and has a TLCC-4 or TLCC-5 activity.

As used interchangeably herein, a "TLCC-4 or TLCC-5 activity", "biological activity of TLCC-4 or TLCC-5" or "functional activity of TLCC-4 or TLCC-5", refers to an activity exerted by a TLCC-4 or TLCC-5 polypeptide or nucleic acid molecule on a TLCC-4 or TLCC-5 responsive cell or tissue, or on a TLCC-4 or TLCC-5 polypeptide substrate, as determined *in vivo*, or *in vitro*, according to standard techniques. In one embodiment, a TLCC-4 or TLCC-5 activity is a direct activity, such as an association with a TLCC-4-target molecule or TLCC-5-target molecule. As used herein, a "target molecule" or "binding partner" is a molecule with which a TLCC-4 or TLCC-5 polypeptide binds or interacts in nature, such that TLCC-4-mediated or TLCC-5-mediated function is achieved. A TLCC-4 or TLCC-5 target molecule can be a non-TLCC-4 or non-TLCC-5 molecule or a TLCC-4 or TLCC-5 polypeptide or polypeptide of the present invention. In an exemplary embodiment, a TLCC-4 or TLCC-5 target molecule is a TLCC-4 or TLCC-5 ligand, *e.g.*, a calcium channel ligand such as calcium. Alternatively, a TLCC-4 or TLCC-5 activity is an indirect activity, such as a cellular signaling activity mediated by interaction of the TLCC-4 or TLCC-5 polypeptide with a TLCC-4 or TLCC-5 ligand. The biological activities of TLCC-4 or TLCC-5 are described herein. For example, the TLCC-4 or TLCC-5 polypeptides of the present invention can have one or more of the following activities: (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, and (7) participate in nociception.

Accordingly, another embodiment of the invention features isolated TLCC-4 or TLCC-5 polypeptides and polypeptides having a TLCC-4 or TLCC-5 activity. Preferred polypeptides are TLCC-4 or TLCC-5 polypeptides having at least one or more of the following domains: an ankyrin repeat domain, and/or a transmembrane domain, and/or a

pore domain, and/or a transient receptor domain, and/or an ion transport protein domain and, preferably, a TLCC-4 or TLCC-5 activity.

Additional preferred polypeptides have one or more of the following domains: an ankyrin repeat domain, a transmembrane domain, a pore domain, a transient receptor domain, and/or an ion transport protein domain, and are, preferably, encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6.

The nucleotide sequence of the isolated human TLCC-4 and TLCC-5 cDNA and the predicted amino acid sequences of the human TLCC-4 and TLCC-5 polypeptides are shown in Figures 1 and 2, and in SEQ ID NOs:1 and 2, and 3 and 4, respectively.

Plasmids containing the nucleotide sequences encoding human TLCC-4 and TLCC-5 were deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on \_\_\_\_\_ and assigned Accession Numbers \_\_\_\_\_. These deposits will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. These deposits were made merely as a convenience for those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112.

The human TLCC-4 gene, which is approximately 3321 nucleotides in length, encodes a polypeptide which is approximately 751 amino acid residues in length. The human TLCC-5 gene, which is approximately 3042 nucleotides in length, encodes a polypeptide which is approximately 1013 amino acid residues in length.

Various aspects of the invention are described in further detail in the following subsections:

#### I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode TLCC-4 or TLCC-5 polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify TLCC-4-encoding or TLCC-5-encoding nucleic acid molecules (*e.g.*, TLCC-4 or TLCC-5 mRNA) and fragments for use as PCR primers for the amplification or mutation of TLCC-4 or TLCC-5 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated TLCC-4 or TLCC-5 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, as a hybridization probe, TLCC-4 or TLCC-5 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to TLCC-4 or TLCC-5 nucleotide

sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In one embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:4. The sequence of SEQ ID NO:1 corresponds to the human TLCC-4. This cDNA comprises sequences encoding the human TLCC-4 polypeptide (*i.e.*, "the coding region", from nucleotides 146-2368) as well as 5' untranslated sequences (nucleotides 1-145) and 3' untranslated sequences (nucleotides 2369-4586). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 146-2368, corresponding to SEQ ID NO:3). Accordingly, in another embodiment, the isolated nucleic acid molecule comprises SEQ ID NO:3 and nucleotides 1-145 and 2369-4586 of SEQ ID NO:1. The sequence of SEQ ID NO:4 corresponds to the human TLCC-5. This cDNA comprises sequences encoding the human TLCC-5 polypeptide (*i.e.*, "the coding region", from nucleotides 1-3042). In yet another embodiment, the nucleic acid molecule consists of the nucleotide sequence set forth as SEQ ID NO:1, 3, 4, or 6.

In still another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to the entire length of the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, (*e.g.*, the entire length of the nucleotide sequence), or to the nucleotide sequences (*e.g.*, the entire length of the nucleotide sequence) of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, or to a portion or complement of any of these nucleotide sequences. In one embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100,



1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 or more nucleotides in length and hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule of SEQ ID NO:1, 3, 4, or 6 or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a TLCC-4 or TLCC-5 polypeptide, *e.g.*, a biologically active portion of a TLCC-4 or TLCC-5 polypeptide. The nucleotide sequence

determined from the cloning of the TLCC-4 or TLCC-5 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other TLCC-4 or TLCC-5 family members, as well as TLCC-4 or TLCC-5 homologues from other species. The probe/primer typically comprises substantially purified oligonucleotide.

The probe/primer (*e.g.*, oligonucleotide) typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, 75, 80, 85, 90, 95, or 100 or more consecutive nucleotides of a sense sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, of an anti-sense sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_.

Exemplary probes or primers are at least (or no greater than) 12 or 15, 20 or 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 or more nucleotides in length and/or comprise consecutive nucleotides of an isolated nucleic acid molecule described herein. Also included within the scope of the present invention are probes or primers comprising contiguous or consecutive nucleotides of an isolated nucleic acid molecule described herein, but for the difference of 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 bases within the probe or primer sequence. Probes based on the TLCC-4 or TLCC-5 nucleotide sequences can be used to detect (*e.g.*, specifically detect) transcripts or genomic sequences encoding the same or homologous polypeptides. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. In another embodiment a set of primers is provided, *e.g.*, primers suitable for use in PCR, which can be used to amplify a selected region of a TLCC-4 or TLCC-5 sequence, *e.g.*, a domain, region, site or other sequence described herein. The primers should be at least 5, 10, or 50 base

pairs in length and less than 100, or less than 200, base pairs in length. The primers should be identical, or differ by no greater than 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 bases when compared to a sequence disclosed herein or to the sequence of a naturally occurring variant. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a TLCC-4 or TLCC-5 polypeptide, such as by measuring a level of a TLCC-4-encoding or TLCC-5-encoding nucleic acid in a sample of cells from a subject, *e.g.*, detecting TLCC-4 or TLCC-5 mRNA levels or determining whether a genomic TLCC-4 or TLCC-5 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a TLCC-4 or TLCC-5 polypeptide" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, which encodes a polypeptide having a TLCC-4 or TLCC-5 biological activity (the biological activities of the TLCC-4 or TLCC-5 polypeptides are described herein), expressing the encoded portion of the TLCC-4 or TLCC-5 polypeptide (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the TLCC-4 or TLCC-5 polypeptide. In an exemplary embodiment, the nucleic acid molecule is at least 50-100, 100-250, 250-500, 500-700, 700-1000, 1000-1250, 1250-1500, 1500-1750, 1750-2000, 2000-2250, 2250-2500, 2500-2750, 2750-300 or more nucleotides in length and encodes a protein having a TLCC-4 or TLCC-5 activity (as described herein).

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_. Such differences can be due to degeneracy of the genetic code, thus resulting in a nucleic acid which encodes the same TLCC-4 or TLCC-5 polypeptides as those encoded by the nucleotide sequence shown in SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a polypeptide having an amino acid sequence which differs by at least 1, but no greater than 5, 10, 20, 50 or 100 amino acid residues from the amino acid sequence shown in SEQ ID NO:2 or 5, or the amino acid sequences encoded by the DNA inserts of the plasmids deposited with the ATCC as Accession Numbers \_\_\_\_\_. In yet another embodiment, the nucleic acid molecule encodes the amino acid sequence of human TLCC-4 or human TLCC-5. If an alignment is needed for this comparison, the sequences should be aligned for maximum homology. Nucleic acid variants can be naturally occurring, such as allelic variants (same locus), homologues (different locus), and orthologues (different organism) or can be non naturally occurring. Non-naturally occurring variants can be made by mutagenesis

techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared to the encoded product).

Allelic variants result, for example, from DNA sequence polymorphisms within a population (e.g., the human population) that lead to changes in the amino acid sequences of the TLCC-4 or TLCC-5 polypeptides. Such genetic polymorphisms in the TLCC-4 or TLCC-5 genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a TLCC-4 or TLCC-5 polypeptide, preferably a mammalian TLCC-4 or TLCC-5 polypeptide, and can further include non-coding regulatory sequences, and introns.

Accordingly, in one embodiment, the invention features isolated nucleic acid molecules which encode a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, or amino acid sequences encoded by the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, wherein the nucleic acid molecule hybridizes to a complement of a nucleic acid molecule comprising SEQ ID NO:1, 3, 4, or 6, for example, under stringent hybridization conditions.

Allelic variants of human TLCC-4 or TLCC-5 include both functional and non-functional TLCC-4 or TLCC-5 polypeptides. Functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC-4 or TLCC-5 polypeptide that maintain the ability to bind a TLCC-4 or TLCC-5 ligand or substrate and/or modulate membrane excitability or signal transduction. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2 or 5, or substitution, deletion or insertion of non-critical residues in non-critical regions of the polypeptide.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human TLCC-4 or TLCC-5 polypeptide that do not have the ability to form functional calcium channels or to modulate membrane excitability. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2 or 5, or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues (e.g., non-human orthologues of the human TLCC-4 or TLCC-5 polypeptides). Orthologues of the human TLCC-4 or TLCC-5 polypeptide are polypeptides that are isolated from non-human organisms and possess the same TLCC-4 or TLCC-5 ligand binding and/or modulation

of membrane excitation mechanisms of the human TLCC-4 or TLCC-5 polypeptide. Orthologues of the human TLCC-4 or TLCC-5 polypeptide can readily be identified as comprising an amino acid sequence that is substantially identical to SEQ ID NO:2 or 5.

Moreover, nucleic acid molecules encoding other TLCC-4 or TLCC-5 family members and, thus, which have a nucleotide sequence which differs from the TLCC-4 or TLCC-5 sequences of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_ are intended to be within the scope of the invention. For example, another TLCC-4 or TLCC-5 cDNA can be identified based on the nucleotide sequence of human TLCC-4 or TLCC-5

5 cDNA can be identified based on the nucleotide sequence of human TLCC-4 or TLCC-5. Moreover, nucleic acid molecules encoding TLCC-4 or TLCC-5 polypeptides from different species, and which, thus, have a nucleotide sequence which differs from the TLCC-4 or TLCC-5 sequences of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_ are intended to be within the scope of the invention. For example, a mouse TLCC-4 or TLCC-5 cDNA can be identified based on the nucleotide sequence of a human TLCC-4 or TLCC-5.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC-4 or TLCC-5 cDNAs of the invention can be isolated based on their homology to the TLCC-4 or TLCC-5 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the TLCC-4 or TLCC-5 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the TLCC-4 or TLCC-5 gene.

Orthologues, homologues and allelic variants can be identified using methods known in the art (e.g., by hybridization to an isolated nucleic acid molecule of the present invention, for example, under stringent hybridization conditions). In one embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_. In another embodiment, the nucleic acid is at least 100,

100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-550, 550-600, 600-650, 650-700, 700-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1070, 1070-1100, 1100-1150, 1150-1200, 1200-1250, 1250-1300, 1300-1350, 1350-1400, 1400-1450, 1450-1500, 1500-1550, 1550-1600, 1600-1650, 1650-1700, 1700-1750, 1750-1800, 1800-1850, 1850-1900, 1900-1950, 1950-2000, 2000-2050, 2050-2100, 2100-2150, 2150-2200, 2200-2250, 2250-2300, 2300-2350,

2350-2400, 2400-2450, 2450-2500, 2500-2550, 2550-2600, 2600-2650, 2650-2700, 2700-2750, 2750-2800, 2800-2850, 2850-2900, 2900-2950, 2950-3000, 3000-3050, 3050-3100, 3100-3150, 3150-3200, 3200-3250, 3250-3300, 3300-3350, 3350-3400, 3400-3450, 3450-3500, 3500-3550, 3550-3600, 3600-3650, 3650-3700, 3700-3750, 3750-3800, 3800-3850, 3850-3900, 3900-3950, 3950-4000, 4000-4050, 4050-4100, 4100-4150, 4150-4200, 4200-4250, 4250-4300, 4300-4350, 4350-4400, 4400-4450, 4450-4500 or more nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, e.g., at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1XSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1XSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature ( $T_m$ ) of the hybrid, where  $T_m$  is determined according to the following equations. For hybrids less than 18 base pairs in length,  $T_m(^{\circ}\text{C}) = 2(\# \text{ of A} + \text{T bases}) + 4(\# \text{ of G} + \text{C bases})$ . For hybrids between 18 and 49 base pairs in length,  $T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - (600/\text{N})$ , where N is the number of bases in the hybrid, and  $[\text{Na}^+]$  is the concentration of sodium ions in the hybridization buffer ( $[\text{Na}^+]$  for

1XSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (e.g., BSA or salmon or herring sperm carrier DNA), detergents (e.g., SDS), chelating agents (e.g., EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH<sub>2</sub>PO<sub>4</sub>, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH<sub>2</sub>PO<sub>4</sub>, 1% SDS at 65°C, see, e.g., Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, 3, 4, or 6 and corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural polypeptide).

In addition to naturally-occurring allelic variants of the TLCC-4 or TLCC-5 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, thereby leading to changes in the amino acid sequence of the encoded TLCC-4 or TLCC-5 polypeptides, without altering the functional ability of the TLCC-4 or TLCC-5 polypeptides. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of TLCC-4 or TLCC-5 (e.g., the sequence of SEQ ID NO:2 or 5) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the TLCC-4 or TLCC-5 polypeptides of the present invention, e.g., those present in a transmembrane domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved between the TLCC-4 or TLCC-5 polypeptides of the present invention and other members of the TLCC-4 or TLCC-5 family are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding TLCC-4 or TLCC-5 polypeptides that contain changes in amino acid residues that are not essential for activity. Such TLCC-4 or TLCC-5 polypeptides differ in amino acid sequence from SEQ ID NO:2 or 5, yet retain biological activity. In one

embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a polypeptide, wherein the polypeptide comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2 or 5 (*e.g.*, to the entire length of SEQ ID NO:2 or 5).

An isolated nucleic acid molecule encoding a TLCC-4 or TLCC-5 polypeptide identical to the polypeptide of SEQ ID NO:2 or 5, can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded polypeptide.

Mutations can be introduced into SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis.

15 Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine,

20 histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, tryptophan), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a TLCC-4 or TLCC-5 polypeptide is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a TLCC-4 or TLCC-5 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for TLCC-4 or TLCC-5 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, the encoded polypeptide can be expressed recombinantly and the activity of the polypeptide can be determined.

In a preferred embodiment, a mutant TLCC-4 or TLCC-5 polypeptide can be assayed for the ability to (1) modulate membrane excitability, (2) influence the resting potential of membranes, (3) modulate wave forms and frequencies of action potentials, (4) modulate thresholds of excitation, (5) modulate neurite outgrowth and synaptogenesis, (6) modulate signal transduction, and (7) participate in nociception.

In addition to the nucleic acid molecules encoding TLCC-4 or TLCC-5 polypeptides described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. In an exemplary embodiment, the invention provides an isolated nucleic acid molecule which is antisense to a TLCC-4 or TLCC-5 nucleic acid molecule (*e.g.*, is antisense to the coding strand of a TLCC-4 or TLCC-5 nucleic acid molecule). An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a polypeptide, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can

10 hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire TLCC-4 or TLCC-5 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding TLCC-4 or TLCC-5. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human TLCC-4 or TLCC-5 corresponds to SEQ ID NO:3 or SEQ ID NO:6, respectively). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding TLCC-4 or TLCC-5. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

20 Given the coding strand sequences encoding TLCC-4 or TLCC-5 disclosed herein (*e.g.*, SEQ ID NO:3 or 6), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of TLCC-4 or TLCC-5 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of TLCC-4 or TLCC-5 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of TLCC-4 or TLCC-5 mRNA (*e.g.*, between the -10 and +10 regions of the start site of a gene nucleotide sequence). An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-

bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N<sup>6</sup>-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N<sup>6</sup>-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxycarbonylmethyluracil, 2-methylthio-D-mannosylqueosine, 5'-methoxycarbonylmethyluracil, 5-methoxyuracil, 2-methylthio-N<sup>6</sup>-isopentenyladenine, uracil-5-oxyacetic acid (y), wbutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5'-oxyacetic acid methyl ester, uracil-5-oxyacetic acid (y), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)'w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a TLCC-4 or TLCC-5 polypeptide to thereby inhibit expression of the polypeptide, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautlier *et al.* (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-

methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave TLCC-4 or TLCC-5 mRNA transcripts to thereby inhibit translation of TLCC-4 or TLCC-5 mRNA. A ribozyme having specificity for a TLCC-4-encoding or TLCC-5-encoding nucleic acid can be designed based upon the nucleotide sequence of a TLCC-4 or TLCC-5 cDNA disclosed herein (*i.e.*, SEQ ID NO:1, 3, 4, or 6, or the nucleotide sequences of the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a TLCC-4-encoding or TLCC-5-encoding mRNA (*see, e.g.*, Czech *et al.* U.S. Patent No. 4,987,071, and Czech *et al.* U.S. Patent No. 5,116,742). Alternatively, TLCC-4 or TLCC-5 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. *See, e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, TLCC-4 or TLCC-5 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the TLCC-4 or TLCC-5 (*e.g.*, the TLCC-4 or TLCC-5 promoter and/or enhancers) to form triple helical structures that prevent transcription of the TLCC-4 or TLCC-5 gene in target cells (*see, generally*, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36, and Maher, L.J. (1992) *Bioassays* 14(12):807-15).

In yet another embodiment, the TLCC-4 or TLCC-5 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (*see, e.g.*, Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* *Proc. Natl. Acad. Sci.* 93: 14670-675.

PNAs of TLCC-4 or TLCC-5 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigenic agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of TLCC-4 or TLCC-5 nucleic acid molecules can also be used in the analysis of single base pair mutations in a gene, (e.g., by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (e.g., S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of TLCC-4 or TLCC-5 can be modified, (e.g., to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of TLCC-4 or TLCC-5 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (e.g., RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag. M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterter, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (see, e.g., Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents (see, e.g., Zon (1988) *Pharm. Res.* 5:539-549). To this end, the

oligonucleotide may be conjugated to another molecule, (e.g., a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

Alternatively, the expression characteristics of an endogenous TLCC-4 or TLCC-5 gene within a cell line or microorganism may be modified by inserting a heterologous DNA regulatory element into the genome of a stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous TLCC-4 or TLCC-5 gene. For example, an endogenous TLCC-4 or TLCC-5 gene which is normally "transcriptionally silent", i.e., a TLCC-4 or TLCC-5 gene which is normally not expressed, or is expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, a transcriptionally silent, endogenous TLCC-4 or TLCC-5 gene may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with an endogenous TLCC-4 or TLCC-5 gene, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described, e.g., in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

## II. Isolated TLCC-4 and TLCC-5 Polypeptides and Anti-TLCC-4 and Anti-TLCC-5 Antibodies

One aspect of the invention pertains to isolated or recombinant TLCC-4 and TLCC-5 polypeptides, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-TLCC-4 or anti-TLCC-5 antibodies. In one embodiment, native TLCC-4 or TLCC-5 polypeptides can be isolated from cells or tissue sources by an appropriate purification scheme using standard polypeptide purification techniques. In another embodiment, TLCC-4 or TLCC-5 polypeptides are produced by recombinant DNA techniques. Alternative to recombinant expression, a TLCC-4 or TLCC-5 polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the TLCC-4 or TLCC-5 polypeptide is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of TLCC-4 or TLCC-5 polypeptide in which the polypeptide is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one

embodiment, the language "substantially free of cellular material" includes preparations of TLCC-4 or TLCC-5 polypeptide having less than about 30% (by dry weight) of non-TLCC-4 or non-TLCC-5 polypeptide (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-TLCC-4 or non-TLCC-5 polypeptide, still more preferably less than about 10% of non-TLCC-4 or non-TLCC-5 polypeptide, and most preferably less than about 5% non-TLCC-4 or non-TLCC-5 polypeptide. When the TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the polypeptide preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC-4 or TLCC-5 polypeptide in which the polypeptide is separated from chemical precursors or other chemicals which are involved in the synthesis of the polypeptide. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of TLCC-4 or TLCC-5 polypeptide having less than about 30% (by dry weight) of chemical precursors or non-TLCC-4 or non-TLCC-5 chemicals, more preferably less than about 20% chemical precursors or non-TLCC-4 or non-TLCC-5 chemicals, still more preferably less than about 10% chemical precursors or non-TLCC-4 or non-TLCC-5 chemicals, and most preferably less than about 5% chemical precursors or non-TLCC-4 or non-TLCC-5 chemicals.

As used herein, a "biologically active portion" of a TLCC-4 or TLCC-5 polypeptide includes a fragment of a TLCC-4 or TLCC-5 polypeptide which participates in an interaction between a TLCC-4 or TLCC-5 molecule and a non-TLCC-4 or non-TLCC-5 molecule. Biologically active portions of a TLCC-4 or TLCC-5 polypeptide include peptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the TLCC-4 or TLCC-5 polypeptide, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 or 5, which include less amino acids than the full length TLCC-4 or TLCC-5 polypeptides, and exhibit at least one activity of a TLCC-4 or TLCC-5 polypeptide. Typically, biologically active portions comprise a domain or motif with at least one activity of the TLCC-4 or TLCC-5 polypeptide, *e.g.*, modulating membrane excitation mechanisms. A biologically active portion of a TLCC-4 or TLCC-5 polypeptide can be a polypeptide which is, for example, 25, 30, 35, 40, 45, 50, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000 or more amino acids in length. Biologically active portions of a TLCC-4

or TLCC-5 polypeptide can be used as targets for developing agents which modulate a TLCC-4 or TLCC-5 mediated activity, *e.g.*, a membrane excitation mechanism.

In one embodiment, a biologically active portion of a TLCC-4 or TLCC-5 polypeptide comprises at least one transmembrane domain. It is to be understood that a preferred biologically active portion of a TLCC-4 or TLCC-5 polypeptide of the present invention comprises at least one or more of the following domains: an ankyrin repeat domain, a transmembrane domain, a pore domain, a transient receptor domain, and/or an ion transport protein domain. Moreover, other biologically active portions, in which other regions of the polypeptide are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native TLCC-4 or TLCC-5 polypeptide.

Another aspect of the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2 or 5, for example, for use as immunogens. In one embodiment, a fragment comprises at least 5 amino acids (*e.g.*, contiguous or consecutive amino acids) of the amino acid sequence of SEQ ID NO:2 or 5, or amino acid sequences encoded by the DNA inserts of the plasmids deposited with the ATCC as Accession Numbers \_\_\_\_\_. In another embodiment, a fragment comprises at least 10, 15, 20, 25, 30, 35, 40, 45, 50 or more amino acids (*e.g.*, contiguous or consecutive amino acids) of the amino acid sequence of SEQ ID NO:2 or 5, or amino acid sequences encoded by the DNA inserts of the plasmids deposited with the ATCC as Accession Numbers \_\_\_\_\_.

In a preferred embodiment, a TLCC-4 or TLCC-5 polypeptide has an amino acid sequence shown in SEQ ID NO:2 or 5. In other embodiments, the TLCC-4 or TLCC-5 polypeptide is substantially identical to SEQ ID NO:2 or 5, and retains the functional activity of the polypeptide of SEQ ID NO:2 or 5, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. In another embodiment, the TLCC-4 or TLCC-5 polypeptide is a polypeptide which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to SEQ ID NO:2 or 5. In another embodiment, the invention features a TLCC-4 or TLCC-5 polypeptide which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to a nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or a complement thereof. This invention further features a TLCC-4 or TLCC-5 polypeptide which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or a complement thereof.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (e.g., when aligning a second sequence to the TLCC-4 amino acid sequence of SEQ ID NO:2 having 742 amino acid residues, at least 222, preferably at least 296, more preferably at least 371, more preferably at least 445, even more preferably at least 519, and even more preferably at least 593 or 667 or more amino acid residues are aligned; when aligning a second sequence to the TLCC-5 amino acid sequence of SEQ ID NO:3 having 1013 amino acid residues, at least 304, preferably at least 405, more preferably at least 506, more preferably at least 608, even more preferably at least 709, and even more preferably at least 810 or 912 or more amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at the GCG website), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at the GCG website), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A preferred, non-limiting example of parameters to be used in conjunction with the GAP program include a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (*Comput. Appl. Biosci.*, 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0 or version 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and polypeptide sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to TLCC-4 or TLCC-5 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 100, wordlength = 3, and a Blossum62 matrix to obtain amino acid sequences homologous to TLCC-4 or TLCC-5 polypeptide molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See the NIH website.

The invention also provides TLCC-4 or TLCC-5 chimeric or fusion proteins. As used herein, a TLCC-4 or TLCC-5 "chimeric protein" or "fusion protein" comprises a TLCC-4 or TLCC-5 polypeptide operatively linked to a non-TLCC-4 or non-TLCC-5 polypeptide. A "TLCC-4 or TLCC-5 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to TLCC-4 or TLCC-5, whereas a "non-TLCC-4 or non-TLCC-5 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the TLCC-4 or TLCC-5 polypeptide, e.g., a protein which is different from the TLCC-4 or TLCC-5 polypeptide and which is derived from the same or a different organism. Within a TLCC-4 or TLCC-5 fusion protein the TLCC-4 or TLCC-5 polypeptide can correspond to all or a portion of a TLCC-4 or TLCC-5 polypeptide. In a preferred embodiment, a TLCC-4 or TLCC-5 fusion protein comprises at least one biologically active portion of a TLCC-4 or TLCC-5 polypeptide. In another preferred embodiment, a TLCC-4 or TLCC-5 fusion protein comprises at least two biologically active portions of a TLCC-4 or TLCC-5 polypeptide. Within the fusion protein, the term "operatively linked" is intended to indicate that the TLCC-4 or TLCC-5 polypeptide and the non-TLCC-4 or non-TLCC-5 polypeptide are fused in-frame to each other. The non-TLCC-4 or non-TLCC-5 polypeptide can be fused to the N-terminus or C-terminus of the TLCC-4 or TLCC-5 polypeptide.



For example, in one embodiment, the fusion protein is a GST-TLCC-4 or GST-TLCC-5 fusion protein in which the TLCC-4 or TLCC-5 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant TLCC-4 or TLCC-5.

In another embodiment, the fusion protein is a TLCC-4 or TLCC-5 polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of TLCC-4 or TLCC-5 can be increased through the use of a heterologous signal sequence.

The TLCC-4 or TLCC-5 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The TLCC-4 or TLCC-5 fusion proteins can be used to affect the bioavailability of a TLCC-4 or TLCC-5 substrate. Use of TLCC-4 or TLCC-5 fusion proteins may be useful therapeutically for the treatment of disorders caused by, for example, (i) aberrant modification or mutation of a gene encoding a TLCC-4 or TLCC-5 polypeptide; (ii) mis-regulation of the TLCC-4 or TLCC-5 gene; and (iii) aberrant post-translational modification of a TLCC-4 or TLCC-5 polypeptide.

Moreover, the TLCC-4 fusion proteins or TLCC-5 fusion proteins of the invention can be used as immunogens to produce anti-TLCC-4 or anti-TLCC-5 antibodies in a subject, to purify TLCC-4 or TLCC-5 ligands and in screening assays to identify molecules which inhibit the interaction of TLCC-4 or TLCC-5 with a TLCC-4 or TLCC-5 substrate.

Preferably, a TLCC-4 or TLCC-5 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.*, John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A TLCC-4-encoding or TLCC-5-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the TLCC-4 or TLCC-5 polypeptide.

The present invention also pertains to variants of the TLCC-4 or TLCC-5 polypeptides which function as either TLCC-4 or TLCC-5 agonists (mimetics) or as TLCC-4 or TLCC-5 antagonists. Variants of the TLCC-4 or TLCC-5 polypeptides can be generated by mutagenesis, e.g., discrete point mutation or truncation of a TLCC-4 or TLCC-5 polypeptide. An agonist of the TLCC-4 or TLCC-5 polypeptides can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a TLCC-4 or TLCC-5 polypeptide. An antagonist of a TLCC-4 or TLCC-5 polypeptide can inhibit one or more of the activities of the naturally occurring form of the TLCC-4 or TLCC-5 polypeptide by, for example, competitively modulating a TLCC-4-mediated activity or TLCC-5-mediated activity of a TLCC-4 or TLCC-5 polypeptide. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the polypeptide has fewer side effects in a subject relative to treatment with the naturally occurring form of the TLCC-4 or TLCC-5 polypeptide.

In one embodiment, variants of a TLCC-4 or TLCC-5 polypeptide which function as either TLCC-4 or TLCC-5 agonists (mimetics) or as TLCC-4 or TLCC-5 antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of a TLCC-4 or TLCC-5 polypeptide for TLCC-4 or TLCC-5 polypeptide agonist or antagonist activity. In one embodiment, a variegated library of TLCC-4 or TLCC-5 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of TLCC-4 or TLCC-5 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential TLCC-4 or TLCC-5 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of TLCC-4 or TLCC-5 sequences therein. There are a variety of methods which can be used to produce libraries of potential TLCC-4 or TLCC-5 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential TLCC-4 or TLCC-5 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) *Terrahedron* 39:3; Iwakura *et al.* (1984) *Ann. Rev. Biochem.* 53:323; Iwakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of a TLCC-4 or TLCC-5 polypeptide coding sequence can be used to generate a variegated population of TLCC-4 or TLCC-5.

fragments for screening and subsequent selection of variants of a TLCC-4 or TLCC-5 polypeptide. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a TLCC-4 or TLCC-5 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the TLCC-4 or TLCC-5 polypeptide.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of TLCC-4 or TLCC-5 polypeptides. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify TLCC-4 or TLCC-5 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated TLCC-4 or TLCC-5 library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, an endothelial cell line, which ordinarily responds to TLCC-4 or TLCC-5 in a particular TLCC-4 or TLCC-5 substrate-dependent manner. The transfected cells are then contacted with TLCC-4 or TLCC-5 and the effect of expression of the mutant on signaling by the TLCC-4 or TLCC-5 substrate can be detected, *e.g.*, by monitoring intracellular calcium, IP<sub>3</sub>, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-4-regulated or TLCC-5-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the TLCC-4 or TLCC-5 substrate, and the individual clones further characterized.

An isolated TLCC-4 or TLCC-5 polypeptide, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind TLCC-4 or TLCC-5 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length

TLCC-4 or TLCC-5 polypeptide can be used or, alternatively, the invention provides antigenic peptide fragments of TLCC-4 or TLCC-5 for use as immunogens. The antigenic peptide of TLCC-4 or TLCC-5 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 or 5 and encompasses an epitope of TLCC-4 or TLCC-5 such that an antibody raised against the peptide forms a specific immune complex with TLCC-4 or TLCC-5. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of TLCC-4 or TLCC-5 that are located on the surface of the polypeptide, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figures 4 and 7).

A TLCC-4 or TLCC-5 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed TLCC-4 or TLCC-5 polypeptide or a chemically synthesized TLCC-4 or TLCC-5 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent.

Immunization of a suitable subject with an immunogenic TLCC-4 or TLCC-5 preparation induces a polyclonal anti-TLCC-4 or anti-TLCC-5 antibody response.

Accordingly, another aspect of the invention pertains to anti-TLCC-4 or anti-TLCC-5 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as TLCC-4 or TLCC-5. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and F(ab')<sub>2</sub> fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind TLCC-4 or TLCC-5. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of TLCC-4 or TLCC-5. A monoclonal antibody composition thus typically displays a single binding affinity for a particular TLCC-4 or TLCC-5 polypeptide with which it immunoreacts.

Polyclonal anti-TLCC-4 or anti-TLCC-5 antibodies can be prepared as described above by immunizing a suitable subject with a TLCC-4 or TLCC-5 immunogen. The anti-TLCC-4 or anti-TLCC-5 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized TLCC-4 or TLCC-5. If desired, the antibody molecules

directed against TLCC-4 or TLCC-5 can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the anti-TLCC-4 or anti-TLCC-5 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497 (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see, generally, R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gelfer *et al.* (1977) *Somatic Cell Gene.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a TLCC-4 or TLCC-5 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds TLCC-4 or TLCC-5.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-TLCC-4 or anti-TLCC-5 monoclonal antibody (see, e.g., G. Galfré *et al.* (1977) *Nature* 266:55052; Gelfer *et al.* *Somatic Cell Gene.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (e.g., a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, e.g., the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which

kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind TLCC-4 or TLCC-5, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-TLCC-4 or anti-TLCC-5 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with TLCC-4 or TLCC-5 to thereby isolate immunoglobulin library members that bind TLCC-4 or TLCC-5. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *Swift/24P™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dover *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breiling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *BioTechnology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Grifflths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrard *et al.* (1991) *BioTechnology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbás *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al.* *Nature* (1990) 348:552-554.

Additionally, recombinant anti-TLCC-4 or anti-TLCC-5 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cahilly *et al.* U.S. Patent No. 4,816,567; Cahilly *et al.* European Patent Application 125,023; Better *et al.* (1988)

Science 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-TLCC-4 or anti-TLCC-5 antibody (*e.g.*, monoclonal antibody) can be used to isolate TLCC-4 or TLCC-5 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-TLCC-4 or anti-TLCC-5 antibody can facilitate the purification of natural TLCC-4 or TLCC-5 from cells and of recombinantly produced TLCC-4 or TLCC-5 expressed in host cells. Moreover, an anti-TLCC-4 or anti-TLCC-5 antibody can be used to detect TLCC-4 or TLCC-5 polypeptide (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the TLCC-4 or TLCC-5 polypeptide. Anti-TLCC-4 or anti-TLCC-5 antibodies can be used diagnostically to monitor polypeptide levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase,  $\beta$ -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{35}\text{S}$  or  $^3\text{H}$ .

### III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, for example recombinant expression vectors, containing a TLCC-4 or TLCC-5 nucleic acid molecule or vectors containing a nucleic acid molecule which encodes a TLCC-4 or TLCC-5 polypeptide (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector,

wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, TLCC-4 or TLCC-5 polypeptides, mutant forms of TLCC-4 or TLCC-5 polypeptides, fusion proteins, and the like).

Accordingly, an exemplary embodiment provides a method for producing a protein, preferably a TLCC-4 or TLCC-5 polypeptide, by culturing in a suitable medium a host cell of the invention (e.g., a mammalian host cell such as a non-human mammalian cell) containing a recombinant expression vector, such that the protein is produced.

The recombinant expression vectors of the invention can be designed for expression of TLCC-4 or TLCC-5 polypeptides in prokaryotic or eukaryotic cells. For example, TLCC-4 or TLCC-5 polypeptides can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRLT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. Purified fusion proteins can be utilized in TLCC-4 or TLCC-5 activity assays, (e.g., direct assays or competitive assays described in detail below), or to generate antibodies specific for TLCC-4 or TLCC-5 polypeptides, for example. In a preferred embodiment, a TLCC-4 or TLCC-5 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (e.g., six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego,

California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118).

Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the TLCC-4 or TLCC-5 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYopSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMfa (Kujian and Herskowitz, (1982) *Cell* 30:933-943), pRKY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picz2 (Invitrogen Corp, San Diego, CA). Alternatively, TLCC-4 or TLCC-5 polypeptides can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith *et al.*, (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd. ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the  $\alpha$ -fetoprotein promoter (Campos and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to TLCC-4 or TLCC-5 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a TLCC-4 or TLCC-5 nucleic acid molecule of the invention is introduced, e.g., a TLCC-4 or TLCC-5 nucleic acid molecule within a vector (e.g., a recombinant expression vector) or a TLCC-4 or TLCC-5 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that

such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a TLCC-4 or TLCC-5 polypeptide can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a TLCC-4 or TLCC-5 polypeptide, or a nucleic acid encoding a selectable marker can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) a TLCC-4 or TLCC-5 polypeptide. Accordingly, the invention further provides methods for producing a TLCC-4 or TLCC-5 polypeptide using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a TLCC-4 or TLCC-5 polypeptide has been introduced) in a suitable medium such that a TLCC-4 or TLCC-5 polypeptide is produced. In another

embodiment, the method further comprises isolating a TLCC-4 or TLCC-5 polypeptide from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which TLCC-4-coding sequences or TLCC-5-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous TLCC-4 or TLCC-5 sequences have been introduced into their genome or homologous recombinant animals in which endogenous TLCC-4 or TLCC-5 sequences have been altered. Such animals are useful for studying the function and/or activity of a TLCC-4 or TLCC-5 and for identifying and/or evaluating modulators of TLCC-4 or TLCC-5 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous TLCC-4 or TLCC-5 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a TLCC-4-encoding nucleic acid or TLCC-5-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The TLCC-4 or TLCC-5 cDNA sequence of SEQ ID NO:1 or 4 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human TLCC-4 or TLCC-5 gene, such as a mouse or rat TLCC-4 or TLCC-5 gene, can be used as a transgene. Alternatively, a TLCC-4 or TLCC-5 gene homologue, such as another TLCC-4 or TLCC-5 family member, can be isolated based on hybridization to the TLCC-4 or TLCC-5 cDNA sequences of SEQ ID NO:1, 3, 4, or 6, or the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_ (described further in subsection I above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a TLCC-4 or TLCC-5 transgene to direct expression of a TLCC-4 or TLCC-5 polypeptide

to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a TLCC-4 or TLCC-5 transgene in its genome and/or expression of TLCC-4 or TLCC-5 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a TLCC-4 or TLCC-5 polypeptide can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a TLCC-4 or TLCC-5 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the TLCC-4 or TLCC-5 gene. The TLCC-4 or TLCC-5 gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3 or 6), but more preferably, is a non-human homologue of a human TLCC-4 or TLCC-5 gene (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1 or 4). For example, a mouse TLCC-4 or TLCC-5 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous TLCC-4 or TLCC-5 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous TLCC-4 or TLCC-5 gene is functionally disrupted (*i.e.*, no longer encodes a functional polypeptide; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous TLCC-4 or TLCC-5 gene is mutated or otherwise altered but still encodes functional polypeptide (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous TLCC-4 or TLCC-5 polypeptide). In the homologous recombination nucleic acid molecule, the altered portion of the TLCC-4 or TLCC-5 gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the TLCC-4 or TLCC-5 gene to allow for homologous recombination to occur between the exogenous TLCC-4 or TLCC-5 gene carried by the homologous recombination nucleic acid molecule and an endogenous TLCC-4 or TLCC-5 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking TLCC-4 or TLCC-5 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid

molecule (see, e.g., Thomas, K.R. and Capecechi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, e.g., an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced TLCC-4 or TLCC-5 gene has

5 homologously recombined with the endogenous TLCC-4 or TLCC-5 gene are selected (see e.g., Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then be injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (see e.g., Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J.

10 Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, e.g., vectors, or homologous 15 recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-human animals can be produced which 20 contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, e.g., Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a 30 recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and 35 induced to exit the growth cycle and enter G<sub>0</sub> phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to

pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, e.g., the somatic cell, is isolated.

#### IV. Pharmaceutical Compositions

5 The TLCC-4 and TLCC-5 nucleic acid molecules, fragments of TLCC-4 and TLCC-5 polypeptides, and anti-TLCC-4 and anti-TLCC-5 antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, polypeptide, or antibody and a pharmaceutically acceptable 10 carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or 15 agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include 20 parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; 25 antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be 30 enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the 35 extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Crenophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage



and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a fragment of a TLCC-4 or TLCC-5 polypeptide or an anti-TLCC-4 or anti-TLCC-5 antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Steroles; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin, or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid.

Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and

therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

As defined herein, a therapeutically effective amount of polypeptide (*i.e.*, an effective dosage) ranges from about 0.001 to 30 mg/kg body weight, preferably about 0.01 to 25 mg/kg body weight, more preferably about 0.1 to 20 mg/kg body weight, and even more preferably about 1 to 10 mg/kg, 2 to 9 mg/kg, 3 to 8 mg/kg, 4 to 7 mg/kg, or 5 to 6 mg/kg body weight. The skilled artisan will appreciate that certain factors may influence the dosage required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a polypeptide or antibody can include a single treatment or, preferably, can include a series of treatments.

In a preferred example, a subject is treated with antibody or polypeptide in the range of between about 0.1 to 20 mg/kg body weight, one time per week for between about 1 to 10 weeks, preferably between about 2 to 8 weeks, more preferably between about 3 to 7 weeks, and even more preferably for about 4, 5, or 6 weeks. It will also be appreciated that the effective dosage of antibody or polypeptide used for treatment may increase or decrease over the course of a particular treatment. Changes in dosage may result and become apparent from the results of diagnostic assays as described herein.

The present invention encompasses agents which modulate expression or activity. An agent may, for example, be a small molecule. For example, such small molecules include, but are not limited to, peptides, peptidomimetics, amino acids, amino acid analogs, polynucleotides, polynucleotide analogs, nucleotides, nucleotide analogs,

organic or inorganic compounds (*i.e.*, including heteroorganic and organometallic compounds) having a molecular weight less than about 10,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 5,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 1,000 grams per mole, organic or inorganic compounds having a molecular weight less than about 500 grams per mole, and salts, esters, and other pharmaceutically acceptable forms of such compounds. It is understood that appropriate doses of small molecule agents depends upon a number of factors within the ken of the ordinarily skilled physician, veterinarian, or researcher. The dose(s) of the small molecule will vary, for example, depending upon the identity, size, and condition of the subject or sample being treated, further depending upon the route by which the composition is to be administered, if applicable, and the effect which the practitioner desires the small molecule to have upon the nucleic acid or polypeptide of the invention.

Exemplary doses include milligram or microgram amounts of the small molecule per kilogram of subject or sample weight (*e.g.*, about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram.

It is

furthermore understood that appropriate doses of a small molecule depend upon the potency of the small molecule with respect to the expression or activity to be modulated. Such appropriate doses may be determined using the assays described herein. When one or more of these small molecules is to be administered to an animal (*e.g.*, a human) in order to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian, or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular animal subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

Further, an antibody (or fragment thereof) may be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mitramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs

or homologues thereof. Therapeutic agents include, but are not limited to, antineoplastic agents (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioposa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), antineoplastic agents (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., daidomycin (formerly actinomycin), bleomycin, mitramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a polypeptide possessing a desired biological activity. Such polypeptides may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin, a protein such as tumor necrosis factor, alpha-interferon, beta-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Armon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfield *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera *et al.* (eds.), pp. 473-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.*, 62:119-58 (1982). Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see, e.g., U.S. Patent 5,328,470), or by stereotactic injection (see, e.g., Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include

the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

#### V. Uses and Methods of the Invention

The nucleic acid molecules, polypeptides, polypeptide homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (e.g., therapeutic and prophylactic). As described herein, a TLCC-4 or TLCC-5 polypeptide of the invention has one or more of the following activities: (1) modulates membrane excitability, (2) influences the resting potential of membranes, (3) modulates wave forms and frequencies of action potentials, (4) modulates thresholds of excitation, (5) modulates neurite outgrowth and synaptogenesis, (6) modulates signal transduction, and (7) participates in nociception.

The isolated nucleic acid molecules of the invention can be used, for example, to express a TLCC-4 or TLCC-5 polypeptide (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect TLCC-4 or TLCC-5 mRNA (e.g., in a biological sample) or a genetic alteration in a TLCC-4 or TLCC-5 gene, and to modulate TLCC-4 or TLCC-5 activity, as described further below. The TLCC-4 or TLCC-5 polypeptides can be used to treat disorders characterized by insufficient or excessive production of a TLCC-4 or TLCC-5 substrate or production of TLCC-4 or TLCC-5 inhibitors. In addition, the TLCC-4 or TLCC-5 polypeptides can be used to screen for naturally occurring TLCC-4 or TLCC-5 substrates, to screen for drugs or compounds which modulate TLCC-4 or TLCC-5 activity, as well as to treat disorders characterized by insufficient or excessive production of TLCC-4 or TLCC-5 polypeptide or production of TLCC-4 or TLCC-5 polypeptide forms which have decreased, aberrant or unvaried activity compared to TLCC-4 or TLCC-5 wild type polypeptide (e.g., CNS disorders (such as neurodegenerative disorders), pain disorders, or cellular growth, differentiation, or migration disorders). Moreover, the anti-TLCC-4 or anti-TLCC-5 antibodies of the invention can be used to detect and isolate TLCC-4 or TLCC-5 polypeptides, to regulate the bioavailability of TLCC-4 or TLCC-5 polypeptides, and modulate TLCC-4 or TLCC-5 activity.

#### A. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to TLCC-4 or TLCC-5 polypeptides, have a stimulatory or inhibitory effect on, for example, TLCC-4 or TLCC-5 expression or TLCC-4 or TLCC-5 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of TLCC-4 or TLCC-5 substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a TLCC-4 or TLCC-5 polypeptide or polypeptide or biologically active portion thereof.

The screening assays of the present invention include, but are not limited to, screening assays specific for calcium receptors, *e.g.*, measuring intracellular  $Ca^{2+}$  concentrations, and/or assaying responses to the presence of the calcium channel stimulators inositol-1,4,5-trisphosphate ( $IP_3$ ) and/or diacylglycerol (DAG) (see, *e.g.*, Hofmann *et al.* (1999) *Nature* 397:259-263; Okada *et al.* (1999) *J. Biol. Chem.* 274:27359-27370; Hofmann *et al.* (2000) *J. Mol. Med.* 78:14-25). Intracellular  $Ca^{2+}$  levels may be assayed, for example, using the fluorescent dye fura-2 as an indicator. Stimulation by DAG may be assayed, for example, using DAG analogues (*e.g.*, OAG (1-oleoyl-2-acetyl-sn-glycerol) and DOG (1,2-dioctanoyl-sn-glycerol)).  $IP_3$  stimulation may be analyzed, for example, using the PLC inhibitor U-73122 in the presence of ATP, which suppresses  $IP_3$ -dependent  $Ca^{2+}$  release.

The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994) *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate TLCC-4 or TLCC-5 activity is determined. Determining the ability of the test compound to modulate TLCC-4 or TLCC-5 activity can be accomplished by monitoring, for example, intracellular calcium,  $IP_3$ , or diacylglycerol concentration, phosphorylation profile of intracellular proteins, or the activity of a TLCC-4-regulated or TLCC-5-regulated transcription factor. The cell, for example, can be of mammalian origin, *e.g.*, a neuronal cell, skin cell, or a liver cell.

The ability of the test compound to modulate TLCC-4 or TLCC-5 binding to a substrate or to bind to TLCC-4 or TLCC-5 can also be determined. Determining the ability of the test compound to modulate TLCC-4 or TLCC-5 binding to a substrate can be accomplished, for example, by coupling the TLCC-4 or TLCC-5 substrate with a radioisotope or enzymatic label such that binding of the TLCC-4 or TLCC-5 substrate to TLCC-4 or TLCC-5 can be determined by detecting the labeled TLCC-4 or TLCC-5 substrate in a complex. Alternatively, TLCC-4 or TLCC-5 could be coupled with a radioisotope or enzymatic label to monitor the ability of a test compound to modulate TLCC-4 or TLCC-5 binding to a TLCC-4 or TLCC-5 substrate in a complex.

Determining the ability of the test compound to bind TLCC-4 or TLCC-5 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to TLCC-4 or TLCC-5 can be determined by detecting the labeled TLCC-4 or TLCC-5 compound in a complex. For example,

compounds (*e.g.*, TLCC-4 or TLCC-5 substrates) can be labeled with  $^{125}I$ ,  $^{35}S$ ,  $^{14}C$ , or  $^3H$ , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a TLCC-4 or TLCC-5 substrate) to interact with TLCC-4 or TLCC-5 without the labeling of any of the interactants. For example, a microphysiometer can be

used to detect the interaction of a compound with TLCC-4 or TLCC-5 without the labeling of either the compound or the TLCC-4 or TLCC-5. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (e.g.,

- 5 CytoSensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and TLCC-4 or TLCC-5.

- 10 In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a TLCC-4 or TLCC-5 target molecule (e.g., a TLCC-4 or TLCC-5 substrate) with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the TLCC-4 or TLCC-5 target molecule. Determining the ability of the test compound to modulate the activity of a TLCC-4 or TLCC-5 target molecule can be accomplished, for example, by determining the ability of the TLCC-4 or TLCC-5 polypeptide to bind to or interact with the TLCC-4 or TLCC-5 target molecule.

- 15 Determining the ability of the TLCC-4 or TLCC-5 polypeptide, or a biologically active fragment thereof, to bind to or interact with a TLCC-4 or TLCC-5 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the TLCC-4 or TLCC-5 polypeptide to bind to or interact with a TLCC-4 or TLCC-5 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular  $Ca^{2+}$ , diacylglycerol,  $IP_3$ , and the like), detecting catalytic/enzymatic activity of the target using an appropriate substrate, 25 detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, e.g., luciferase), or detecting a target-regulated cellular response.

- 30 In yet another embodiment, an assay of the present invention is a cell-free assay in which a TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof is determined. Preferred biologically active portions of the TLCC-4 or TLCC-5 polypeptides to be used in assays of the present invention include fragments which participate in interactions with non-TLCC-4 or non-TLCC-5 molecules, e.g., fragments with high surface 35 probability scores (see, for example, Figures 4 and 7). Binding of the test compound to the TLCC-4 or TLCC-5 polypeptide can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof with a known compound

which binds TLCC-4 or TLCC-5 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a TLCC-4 or TLCC-5 polypeptide, wherein determining the ability of the test compound to interact with a TLCC-4 or TLCC-5 polypeptide comprises determining the ability of the test compound to preferentially bind to TLCC-4 or TLCC-5 or biologically active portion thereof as compared to the known compound.

- 5 In another embodiment, the assay is a cell-free assay in which a TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a TLCC-4 or TLCC-5 polypeptide can be accomplished, for example, by determining the ability of the TLCC-4 or TLCC-5 polypeptide to bind to a TLCC-4 or TLCC-5 target molecule by one of the methods described above for determining direct binding.

- 10 Determining the ability of the TLCC-4 or TLCC-5 polypeptide to bind to a TLCC-4 or TLCC-5 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjlander, S. and Urbanicky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (e.g., BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

- 20 In an alternative embodiment, determining the ability of the test compound to modulate the activity of a TLCC-4 or TLCC-5 polypeptide can be accomplished by determining the ability of the TLCC-4 or TLCC-5 polypeptide to further modulate the activity of a downstream effector of a TLCC-4 or TLCC-5 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

- 30 In yet another embodiment, the cell-free assay involves contacting a TLCC-4 or TLCC-5 polypeptide or biologically active portion thereof with a known compound which binds the TLCC-4 or TLCC-5 polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the TLCC-4 or TLCC-5 polypeptide, wherein determining the ability of the test compound to interact with the TLCC-4 or TLCC-5 polypeptide comprises determining the ability of the TLCC-4 or TLCC-5 polypeptide to preferentially bind to or modulate the activity of a TLCC-4 or TLCC-5 target molecule. 35

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either TLCC-4 or TLCC-5 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the polypeptides, as well as to accommodate automation of the assay. Binding of a test compound to a TLCC-4 or TLCC-5 polypeptide, or interaction of a TLCC-4 or TLCC-5 polypeptide with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/TLCC-4 or TLCC-5 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized micrometer plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or TLCC-4 or TLCC-5 polypeptide, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or micrometer plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of TLCC-4 or TLCC-5 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a TLCC-4 or TLCC-5 polypeptide or a TLCC-4 or TLCC-5 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated TLCC-4 or TLCC-5 polypeptide or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with TLCC-4 or TLCC-5 polypeptide or target molecules but which do not interfere with binding of the TLCC-4 or TLCC-5 polypeptide to its target molecule can be derivatized to the wells of the plate, and unbound target or TLCC-4 or TLCC-5 polypeptide trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the TLCC-4 or TLCC-5 polypeptide or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the TLCC-4 or TLCC-5 polypeptide or target molecule.

In another embodiment, modulators of TLCC-4 or TLCC-5 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of TLCC-4 or TLCC-5 mRNA or polypeptide in the cell is determined. The level of expression of TLCC-4 or TLCC-5 mRNA or polypeptide in the presence of the candidate compound is compared to the level of expression of TLCC-4 or TLCC-5 mRNA or polypeptide in the absence of the candidate compound. The candidate compound can then be identified as a modulator of TLCC-4 or TLCC-5 expression based on this comparison. For example, when expression of TLCC-4 or TLCC-5 mRNA or polypeptide is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of TLCC-4 or TLCC-5 mRNA or polypeptide expression. Alternatively, when expression of TLCC-4 or TLCC-5 mRNA or polypeptide is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of TLCC-4 or TLCC-5 mRNA or polypeptide expression. The level of TLCC-4 or TLCC-5 mRNA or polypeptide expression in the cells can be determined by methods described herein for detecting TLCC-4 or TLCC-5 mRNA or polypeptide.

In yet another aspect of the invention, the TLCC-4 or TLCC-5 polypeptides can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with TLCC-4 or TLCC-5 ("TLCC-4-binding proteins or TLCC-5-binding proteins" or "TLCC-4 bp or TLCC-5 bp") and are involved in TLCC-4 or TLCC-5 activity. Such TLCC-4-binding proteins or TLCC-5-binding proteins are also likely to be involved in the propagation of signals by the TLCC-4 or TLCC-5 polypeptides or TLCC-4 or TLCC-5 targets as, for example, downstream elements of a TLCC-4-mediated or TLCC-5-mediated signaling pathway. Alternatively, such TLCC-4-binding proteins or TLCC-5-binding proteins are likely to be TLCC-4 or TLCC-5 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a TLCC-4 or TLCC-5 polypeptide is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*,

forming a TLCC-4-dependent or TLCC-5-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the TLCC-4 or TLCC-5 polypeptide.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a TLCC-4 or TLCC-5 polypeptide can be confirmed *in vivo*, e.g., in an animal such as an animal model for cellular transformation and/or tumorigenesis.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (e.g., a TLCC-4 or TLCC-5 modulating agent, an antisense TLCC-4 or TLCC-5 nucleic acid molecule, a TLCC-4-specific or TLCC-5-specific antibody, or a TLCC-4-binding partner or TLCC-5-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

Models for studying pain *in vivo* include rat models of neuropathic pain caused by methods such as intraperitoneal administration of Taxol (Auhier *et al.* (2000) *Brain Res.* 887:239-249), chronic constriction injury (CCI), partial sciatic nerve transection (Linenlaub and Sommer (2000) *Pain* 89:97-106), transection of the tibial and sural nerves (Lee *et al.* (2000) *Neurosci. Lett.* 291:29-32), the spared nerve injury model (Decosterd and Woolf (2000) *Pain* 87:149-158), cuffing the sciatic nerve (Pritchett and Henry (2000) *Eur. J. Neurosci.* 12:2006-2020), unilateral tight ligation (Esser and Sawynok (2000) *Eur. J. Pharmacol.* 399:131-139), L5 spinal nerve ligation (Honore *et al.* (2000) *Neurosci.* 98:585-598), and photochemically induced ischemic nerve injury (Hao *et al.* (2000) *Exp. Neurol.* 163:231-238); rat models of nociceptive pain caused by methods such as the Chung Method, the Bennett Method, and intraperitoneal administration of Complete Freund's adjuvant (Abdi *et al.* (2000) *Anesth. Analg.* 91:955-959); rat models of post-incisional pain caused by incising the skin and fascia of a hind paw (Olivera and Prado (2000) *Brax. J. Med. Res.* 33:957-960); rat models of cancer pain caused by methods such as injecting osteolytic sarcoma cells into the

femur (Honore *et al.* (2000) *Neurosci.* 98:585-598); and rat models of visceral pain caused by methods such as intraperitoneal administration of cyclophosphamide.

Various methods of determining an animal's response to pain are known in the art. Examples of such methods include, but are not limited to brief intense exposure to a focused heat source, administration of a noxious chemical subcutaneously, the tail flick test, the hot plate test, the formalin test, the Von Frey threshold test, and testing for stress-induced analgesia (e.g., by restraint, foot shock, and/or cold water swim) (Crawley (2000) *What's Wrong With My Mouse?* Wiley-Liss pp. 72-75).

#### 10 B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

##### 20 1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the TLCC-4 or TLCC-5 nucleotide sequences, described herein, can be used to map the location of the TLCC-4 or TLCC-5 genes on a chromosome. The mapping of the TLCC-4 or TLCC-5 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, TLCC-4 or TLCC-5 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the TLCC-4 or TLCC-5 nucleotide sequences. Computer analysis of the TLCC-4 or TLCC-5 sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the TLCC-4 or TLCC-5 sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they

lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the TLCC-4 or TLCC-5 nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a TLCC-4 or TLCC-5 sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, *Human Chromosomes: A Manual of Basic Techniques* (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, *Mendelian Inheritance in Man*, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature*, 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the TLCC-4 or TLCC-5 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

## 2. Tissue Typing

The TLCC-4 or TLCC-5 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the TLCC-4 or TLCC-5 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals



and from tissue. The TLCC-4 or TLCC-5 nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 or 4 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 or 6 are used, a more appropriate number of primers for positive individual identification would be 500-2,000. If a panel of reagents from TLCC-4 or TLCC-5 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

### 20 3. Use of TLCC-4 and TLCC-5 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology.

Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.*, another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 or 4 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the TLCC-4 or TLCC-5 nucleotide

sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 or 4 having a length of at least 20 bases, preferably at least 30 bases.

The TLCC-4 or TLCC-5 nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such TLCC-4 or TLCC-5 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, TLCC-4 or TLCC-5 primers or probes can be used to screen tissue culture for contamination (*i.e.*, screen for the presence of a mixture of different types of cells in a culture).

### C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically.

Accordingly, one aspect of the present invention relates to diagnostic assays for determining TLCC-4 or TLCC-5 polypeptide and/or nucleic acid expression as well as TLCC-4 or TLCC-5 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with TLCC-4 or TLCC-5 polypeptide, nucleic acid expression or activity. For example, mutations in a TLCC-4 or TLCC-5 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with TLCC-4 or TLCC-5 polypeptide, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of TLCC-4 or TLCC-5 in clinical trials.

These and other agents are described in further detail in the following sections.

### 35 1. Diagnostic Assays

An exemplary method for detecting the presence or absence of TLCC-4 or TLCC-5 polypeptide or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a

compound or an agent capable of detecting TLCC-4 or TLCC-5 polypeptide or nucleic acid (e.g., mRNA, or genomic DNA) that encodes TLCC-4 or TLCC-5 polypeptide such that the presence of TLCC-4 or TLCC-5 polypeptide or nucleic acid is detected in the biological sample. In another aspect, the present invention provides a method for detecting the presence of TLCC-4 or TLCC-5 activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of TLCC-4 or TLCC-5 activity such that the presence of TLCC-4 or TLCC-5 activity is detected in the biological sample. A preferred agent for detecting TLCC-4 or TLCC-5 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to TLCC-4 or TLCC-5 mRNA or genomic DNA. The nucleic acid probe can be, for example, the TLCC-4 or TLCC-5 nucleic acid set forth in SEQ ID NO:1, 3, 4, or 6, or the DNA inserts of the plasmids deposited with ATCC as Accession Numbers \_\_\_\_\_, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to TLCC-4 or TLCC-5 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting TLCC-4 or TLCC-5 polypeptide is an antibody capable of binding to TLCC-4 or TLCC-5 polypeptide, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or F(ab')<sub>2</sub>) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect TLCC-4 or TLCC-5 mRNA, polypeptide, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of TLCC-4 or TLCC-5 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of TLCC-4 or TLCC-5 polypeptide include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of TLCC-4 or TLCC-5 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of TLCC-4 or TLCC-5 polypeptide include introducing into a subject a labeled anti-TLCC-4 or anti-TLCC-5 antibody. For example, the antibody can be labeled with a radioactive

marker whose presence and location in a subject can be detected by standard imaging techniques.

The present invention also provides diagnostic assays for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a TLCC-4 or TLCC-5 polypeptide; (ii) aberrant expression of a gene encoding a TLCC-4 or TLCC-5 polypeptide; (iii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a TLCC-4 or TLCC-5 polypeptide, wherein a wild-type form of the gene encodes a protein with a TLCC-4 or TLCC-5 activity. "Misexpression or aberrant expression", as used herein, refers to a non-wild type pattern of gene expression, at the RNA or protein level. It includes, but is not limited to, expression at non-wild type levels (e.g., over or under expression); a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed (e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage); a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-translational modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene (e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus).

In one embodiment, the biological sample contains polypeptide molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting TLCC-4 or TLCC-5 polypeptide, mRNA, or genomic DNA, such that the presence of TLCC-4 or TLCC-5 polypeptide, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of TLCC-4 or TLCC-5 polypeptide, mRNA or genomic DNA in the control sample with the presence of TLCC-4 or TLCC-5 polypeptide, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of TLCC-4 or TLCC-5 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting TLCC-4 or TLCC-5 polypeptide or mRNA in a biological sample; means for determining the amount of TLCC-4 or TLCC-5 in the sample; and

means for comparing the amount of TLCC-4 or TLCC-5 in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect TLCC-4 or TLCC-5 polypeptide or nucleic acid.

## 2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity. As used herein, the term "aberrant" includes a TLCC-4 or TLCC-5 expression or activity which deviates from the wild type TLCC-4 or TLCC-5 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant TLCC-4 or TLCC-5 expression or activity is intended to include the cases in which a mutation in the TLCC-4 or TLCC-5 gene causes the TLCC-4 or TLCC-5 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional TLCC-4 or TLCC-5 polypeptide or a polypeptide which does not function in a wild-type fashion, e.g., a polypeptide which does not interact with a TLCC-4 or TLCC-5 substrate, e.g., a non-calcium channel subunit or ligand and/or a non-vanilloid receptor subunit or ligand, or one which interacts with a non-TLCC-4 or non-TLCC-5 substrate, e.g. a non-calcium channel subunit or ligand and/or a non-vanilloid receptor subunit or ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response, such as cellular proliferation. For example, the term unwanted includes a TLCC-4 or TLCC-5 expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in TLCC-4 or TLCC-5 polypeptide activity or nucleic acid expression, such as a CNS disorder (e.g., a neurodegenerative disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder). Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in TLCC-4 or TLCC-5 polypeptide activity or nucleic acid expression, such as a CNS disorder, a pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity in which a test sample is obtained from a subject and TLCC-4 or TLCC-5 polypeptide or nucleic acid (e.g., mRNA or genomic DNA) is detected, wherein the presence of TLCC-4 or TLCC-5

polypeptide or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity in which a test sample is obtained and TLCC-4 or TLCC-5 polypeptide or nucleic acid expression or activity is detected (e.g., wherein the abundance of TLCC-4 or TLCC-5 polypeptide or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity).

The methods of the invention can also be used to detect genetic alterations in a TLCC-4 or TLCC-5 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in TLCC-4 or TLCC-5 polypeptide activity or nucleic acid expression, such as a CNS disorder, pain disorder, or a disorder of cellular growth, differentiation, or migration. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a TLCC-4 or TLCC-5 - polypeptide, or the mis-expression of the TLCC-4 or TLCC-5 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of the following: 1) a deletion of one or more

nucleotides from a TLCC-4 or TLCC-5 gene; 2) an addition of one or more nucleotides to a TLCC-4 or TLCC-5 gene; 3) a substitution of one or more nucleotides of a TLCC-4 or TLCC-5 gene; 4) a chromosomal rearrangement of a TLCC-4 or TLCC-5 gene; 5) an alteration in the level of a messenger RNA transcript of a TLCC-4 or TLCC-5 gene; 6) aberrant modification of a TLCC-4 or TLCC-5 gene, such as of the methylation pattern of the genomic DNA; 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a TLCC-4 or TLCC-5 gene; 8) a non-wild type level of a TLCC-4 or TLCC-5 polypeptide; 9) allelic loss of a TLCC-4 or TLCC-5 gene; and 10) inappropriate post-translational modification of a TLCC-4 or TLCC-5 polypeptide. As

described herein, there are a large number of assays known in the art which can be used for detecting alterations in a TLCC-4 or TLCC-5 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the TLCC-4 or TLCC-5 gene (see, e.g., Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a TLCC-4 or TLCC-5 gene under conditions such that hybridization and amplification of the TLCC-4 or TLCC-5 gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a TLCC-4 or TLCC-5 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in TLCC-4 or TLCC-5 can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high

density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in TLCC-4 or TLCC-5 can be identified in two dimensional arrays containing light-generated DNA probes as

described in Cronin, M.T. *et al. supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the TLCC-4 or TLCC-5 gene and detect mutations by comparing the sequence of the sample TLCC-4 or TLCC-5 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the TLCC-4 or TLCC-5 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type TLCC-4 or TLCC-5 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digest the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation (see, for example, Cotton *et al.* (1988) *Proc. Natl. Acad. Sci. USA*

85:4397, Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295). In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in TLCC-4 or TLCC-5 cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from *HeLa* cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based

on a TLCC-4 or TLCC-5 sequence, e.g., a wild-type TLCC-4 or TLCC-5 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like (see, for example, U.S. Patent No. 5,459,039).

In other embodiments, alterations in electrophoretic mobility will be used to

identify mutations in TLCC-4 or TLCC-5 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc. Natl. Acad. Sci. USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments

of sample and control TLCC-4 or TLCC-5 nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Kcen *et al.* (1991) *Trends Genet.* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:465). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys. Chem.* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective

primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saito *et al.* (1986) *Nature* 324:163); Saito *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention.

Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a TLCC-4 or TLCC-5 gene.

Furthermore, any cell type or tissue in which TLCC-4 or TLCC-5 is expressed may be utilized in the prognostic assays described herein.

### 3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (e.g., drugs) on the expression or activity of a TLCC-4 or TLCC-5 polypeptide (e.g., the modulation of membrane excitability) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase TLCC-4 or TLCC-5 gene expression, polypeptide levels, or upregulate TLCC-4 or TLCC-5 activity, can be monitored in clinical trials of subjects exhibiting decreased TLCC-4 or TLCC-5 gene expression, polypeptide levels, or downregulated TLCC-4 or TLCC-5 activity. Alternatively, the effectiveness of an agent determined by a screening

assay to decrease TLCC-4 or TLCC-5 gene expression, polypeptide levels, or downregulate TLCC-4 or TLCC-5 activity, can be monitored in clinical trials of subjects exhibiting increased TLCC-4 or TLCC-5 gene expression, polypeptide levels, or upregulated TLCC-4 or TLCC-5 activity. In such clinical trials, the expression or activity of a TLCC-4 or TLCC-5 gene, and preferably, other genes that have been implicated in, for example, a TLCC-4-associated disorder or TLCC-5-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including TLCC-4 or TLCC-5, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates TLCC-4 or TLCC-5 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on TLCC-4-associated disorders or TLCC-5-associated disorders (*e.g.*, disorders characterized by deregulated signaling or membrane excitation), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of TLCC-4 or TLCC-5 and other genes implicated in the TLCC-4-associated disorder or TLCC-5-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of polypeptide produced, by one of the methods as described herein, or by measuring the levels of activity of TLCC-4 or TLCC-5 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a TLCC-4 or TLCC-5 polypeptide, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the TLCC-4 or TLCC-5 polypeptide, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the TLCC-4 or TLCC-5 polypeptide, mRNA, or genomic DNA in the pre-administration sample with the TLCC-4 or TLCC-5 polypeptide, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of TLCC-4 or TLCC-5 to higher levels

than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of TLCC-4 or TLCC-5 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, TLCC-4 or TLCC-5 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

4. Electronic Apparatus Readable Media and Arrays  
Electronic apparatus readable media comprising TLCC-4 or TLCC-5 sequence information is also provided. As used herein, "TLCC-4 or TLCC-5 sequence information" refers to any nucleotide and/or amino acid sequence information particular to the TLCC-4 or TLCC-5 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said TLCC-4 or TLCC-5 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantitative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon TLCC-4 or TLCC-5 sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can

readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the TLCC-4 or TLCC-5 sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of dalaprocessor structuring formats (e.g., text file or database) may be employed in order to obtain or create a medium having recorded thereon the TLCC-4 or TLCC-5 sequence information.

By providing TLCC-4 or TLCC-5 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder, wherein the method comprises the steps of determining TLCC-4 or TLCC-5 sequence information associated with the subject and based on the TLCC-4 or TLCC-5 sequence information, determining whether the subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a disease associated with a TLCC-4 or TLCC-5 wherein the method comprises the steps of determining TLCC-4 or TLCC-5 sequence information associated with the subject, and based on the TLCC-4 or TLCC-5 sequence information, determining whether the subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder associated with TLCC-4 or TLCC-5, said method comprising the steps of receiving TLCC-4 or TLCC-5 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to TLCC-4 or TLCC-5 and/or a TLCC-4 or TLCC-5-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC-4 or TLCC-5 information (e.g., sequence information and/or information related thereto), and the acquired information, determining whether the subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder, said method comprising the steps of receiving information related to TLCC-4 or TLCC-5 (e.g., sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to TLCC-4 or TLCC-5 and/or related to a TLCC-4 or TLCC-5-associated disease or disorder, and based on one or more of the phenotypic information, the TLCC-4 or TLCC-5 information, and the acquired information, determining whether the subject has a TLCC-4 or TLCC-5-associated disease or disorder or a pre-disposition to a TLCC-4 or TLCC-5-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a TLCC-4 or TLCC-5 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be TLCC-4 or TLCC-5. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene

expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a TLCC-4 or TLCC-5-associated disease or disorder, progression of TLCC-4 or TLCC-5-associated disease or disorder, and processes, such as a cellular transformation associated with the TLCC-4 or TLCC-5-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (e.g., ascertaining the effect of TLCC-4 or TLCC-5 expression on the expression of other genes). This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (e.g., including TLCC-4 or TLCC-5) that could serve as a molecular target for diagnosis or therapeutic intervention.

#### D. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted TLCC-4 or TLCC-5 expression or activity, e.g. a CNS disorder, pain disorder, or a cellular proliferation, growth, differentiation, or migration disorder. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers to the study of how a patient's genes determine his or her response to a drug (e.g., a patient's "drug response phenotype", or "drug response

genotype"). Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the TLCC-4 or TLCC-5 molecules of the present invention or TLCC-4 or TLCC-5 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

Treatment is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease, a symptom of disease or a predisposition toward a disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease, the symptoms of disease or the predisposition toward disease.

A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides.

#### 1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted TLCC-4 or TLCC-5 expression or activity, by administering to the subject a TLCC-4 or TLCC-5 or an agent which modulates TLCC-4 or TLCC-5 expression or at least one TLCC-4 or TLCC-5 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted TLCC-4 or TLCC-5 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein.

Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the TLCC-4 or TLCC-5 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of TLCC-4 or TLCC-5 aberrancy, for example, a TLCC-4 or TLCC-5 agonist, or TLCC-4 or TLCC-5 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

#### 2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating TLCC-4 or TLCC-5 expression or activity for therapeutic purposes. Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell capable of expressing TLCC-4 or TLCC-5 with an agent that modulates one or more of the activities of TLCC-4 or TLCC-5 polypeptide activity associated with the cell, such that TLCC-4 or TLCC-5 activity in the cell is modulated. An agent that modulates TLCC-4



or TLCC-5 polypeptide activity can be an agent as described herein, such as a nucleic acid or a polypeptide, a naturally-occurring target molecule of a TLCC-4 or TLCC-5 polypeptide (e.g., a TLCC-4 or TLCC-5 substrate), a TLCC-4 or TLCC-5 antibody, a TLCC-4 or TLCC-5 agonist or antagonist, a peptidomimetic of a TLCC-4 or TLCC-5 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more TLCC-4 or TLCC-5 activities. Examples of such stimulatory agents include active TLCC-4 or TLCC-5 polypeptide and a nucleic acid molecule encoding TLCC-4 or TLCC-5 that has been introduced into the cell. In another embodiment, the agent inhibits one or more TLCC-4 or TLCC-5 activities. Examples of such inhibitory agents include antisense TLCC-4 or TLCC-5 nucleic acid molecules, anti-TLCC-4 or anti-TLCC-5 antibodies, and TLCC-4 or TLCC-5 inhibitors. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a TLCC-4 or TLCC-5 polypeptide or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) TLCC-4 or TLCC-5 expression or activity. In another embodiment, the method involves administering a TLCC-4 or TLCC-5 polypeptide or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted TLCC-4 or TLCC-5 expression or activity.

Stimulation of TLCC-4 or TLCC-5 activity is desirable in situations in which TLCC-4 or TLCC-5 is abnormally downregulated and/or in which increased TLCC-4 or TLCC-5 activity is likely to have a beneficial effect. Likewise, inhibition of TLCC-4 or TLCC-5 activity is desirable in situations in which TLCC-4 or TLCC-5 is abnormally upregulated and/or in which decreased TLCC-4 or TLCC-5 activity is likely to have a beneficial effect.

### 3. Pharmacogenomics

The TLCC-4 or TLCC-5 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on TLCC-4 or TLCC-5 activity (e.g., TLCC-4 or TLCC-5 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) TLCC-4-associated disorders or TLCC-5-associated disorders (e.g., proliferative disorders) associated with aberrant or unwanted TLCC-4 or TLCC-5 activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound

or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a TLCC-4 or TLCC-5 molecule or TLCC-4 or TLCC-5 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a TLCC-4 or TLCC-5 molecule or TLCC-4 or TLCC-5 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11): 983-985 and Linder, M. W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (e.g., a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase I/II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drug target is known (e.g., a TLCC-4 or TLCC-5 polypeptide of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (e.g., a TLCC-4 or TLCC-5 molecule or TLCC-4 or TLCC-5 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a TLCC-4 or TLCC-5 molecule or TLCC-4 or TLCC-5 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

#### 4. Use of TLCC-4 and TLCC-5 Molecules as Surrogate Markers

The TLCC-4 or TLCC-5 molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the presence, absence and/or quantity of the TLCC-4 or TLCC-5 molecules of the invention may be detected, and may be correlated with one or more biological states *in vivo*. For example, the TLCC-4 or TLCC-5 molecules of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (e.g., with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (e.g., early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (e.g., an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The TLCC-4 or TLCC-5 molecules of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection

of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker (e.g., a TLCC-4 or TLCC-5 marker) transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, anti-TLCC-4 or anti-TLCC-5 antibodies may be employed in an immune-based detection system for a TLCC-4 or TLCC-5 protein marker, or TLCC-4-specific or TLCC-5-specific radiolabeled probes may be used to detect a TLCC-4 or TLCC-5 mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schenag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

The TLCC-4 or TLCC-5 molecules of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates with a specific clinical drug response or susceptibility in a subject (see, e.g., McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA, or protein (e.g., TLCC-4 or TLCC-5 protein or RNA) for specific tumor markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation in TLCC-4 or TLCC-5 DNA may correlate TLCC-4 or TLCC-5 drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application, as well as the Figures and the Sequence Listings, are incorporated herein by reference.

## EXAMPLES

### EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN

#### TLCC-4 AND TLCC-5 cDNA

In this example, the identification and characterization of the genes encoding human TLCC-4 (clone Fbh48000) and TLCC-5 (clone Fbh52920c) is described.

#### 10 Isolation of the TLCC-4 and TLCC-5 cDNA

The invention is based, at least in part, on the discovery of human genes encoding novel polypeptides, referred to herein as TLCC-4 and TLCC-5. The entire sequence of the human clone Fbh48000 was determined and found to contain an open reading frame termed human "TLCC-4." The nucleotide sequence of the human TLCC-4 gene is set forth in Figure 1 and in the Sequence Listing as SEQ ID NO:1. The amino acid sequence of the human TLCC-4 expression product is set forth in Figure 1 and in the Sequence Listing as SEQ ID NO:2. The TLCC-4 polypeptide encoded by this nucleic acid comprises about 751 amino acids. Clone Fbh48000, comprising the coding region of human TLCC-4, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on \_\_\_\_\_, and assigned Accession No. \_\_\_\_\_.

The entire sequence of the human clone Fbh52920c was determined and found to contain an open reading frame termed human "TLCC-5." The nucleotide sequence of the human TLCC-5 gene is set forth in Figure 2 and in the Sequence Listing as SEQ ID NO:4. The amino acid sequence of the human TLCC-4 expression product is set forth in Figure 2 and in the Sequence Listing as SEQ ID NO:5. The TLCC-5 polypeptide comprises about 1013 amino acids. Clone Fbh48000, comprising the coding region of human TLCC-5, was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on \_\_\_\_\_, and assigned Accession No. \_\_\_\_\_.

#### Analysis of the Human TLCC-4 and TLCC-5 Molecules

The human TLCC-5 amino acid sequence was aligned with the amino acid sequence of transient receptor potential polypeptide 7 (TRP7) and melastatin from *Homo sapiens* using the CLUSTAL W (1.74) multiple sequence alignment program. The results of the alignments are set forth in Figures 10 and 11.

A search was performed against the HMM database in PFAM (Figure 5) resulting in the identification of four ankyrin repeat domains at about residues 167-202 (score = 1.6), 214-246 (score = 30.6), 261-294 (score = 27.9), and 340-372 (score = 18.6), and an ion transport protein domain at about residues 510-677 (score = 34.5) in the amino acid sequence of human TLCC-4 (SEQ ID NO:2).

A search was also performed against the MEMSAT database resulting in the identification of six transmembrane domains in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 440-461, 488-508, 520-540, 547-565, 590-609, and 652-676 (Figure 6).

A search was further performed against the HMM database resulting in the identification of two transient receptor domains at about residues 720-778 (score = 21.7) and 820-876 (score = 1.5), in the amino acid sequence of human TLCC-5 (SEQ ID NO:5) (Figure 8).

A search was also performed against the MEMSAT database resulting in the identification of two transmembrane domains in the amino acid sequence of human TLCC-5 (SEQ ID NO:5) at about residues 786-803 and 826-848 (Figure 9).

A search in the Prosite database further resulted in the identification of several protein kinase C phosphorylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 37-39, 167-169, 290-292, 335-337, 374-376, 476-478, 498-500, and 688-690; several N-glycosylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 452-455 and 683-686; a cAMP- and cGMP-dependent protein kinase phosphorylation site in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 375-378; several casein kinase II phosphorylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 88-91, 163-166, 290-293, 305-308, 312-315, 388-391, 393-396, 397-400, 402-405, 411-414, 498-501, 607-610, 624-627, and 699-702; several tyrosine kinase phosphorylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 253-260, 375-382, and 614-622; several N-myristoylation sites in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 238-243 and 602-607; an amidation site in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 12-15; and a leucine zipper site in the amino acid sequence of human TLCC-4 (SEQ ID NO:2) at about residues 584-605.

A search performed in the Prosite database further resulted in the identification of several protein kinase C phosphorylation sites in the amino acid sequence of human TLCC-5 (SEQ ID NO:5) at about residues 21-23, 28-30, 39-41, 105-107, 240-242, 305-307, 331-333, 338-340, 711-713, 802-804, 901-903, 972-974, and 1001-1003; several casein kinase II phosphorylation sites in the amino acid sequence of human TLCC-5 (SEQ ID NO:5) at about residues 54-57, 143-146, 223-226, 240-243, 308-311, 360-363,

436-439, 487-490, 576-579, 725-728, 977-980, and 982-985; and several tyrosine kinase phosphorylation sites in the amino acid sequence of human TLCC-5 (SEQ ID NO:5) at about residues 49-55, 247-254, and 307-314.

Further domain motifs were identified by using the amino acid sequence of TLCC-4 (SEQ ID NO:2) to search through the ProDom database. Numerous matches against protein domains described as "receptor vanilloid channel activated receptor-related receptor-like type OTRPC4", "channel vanilloid receptor activated receptor-related receptor-like OTRPC4 2B ion", "repeat ankyrin kinase nuclear factor channel", "ankyrin repeat kinase domain UNC-44 alternative glycoprotein EGF-like", "ankyrin", "channel osmotically receptor-related vanilloid cation", "receptor vanilloid channel activated receptor-related receptor-like calcium type", "calcium epithelial channel transporter homolog CAT2", "channel protein receptor calcium transient potential transmembrane ion transport", and "receptor vanilloid channel activated osmotically", and the like were identified.

Further domain motifs were identified by using the amino acid sequence of TLCC-5 (SEQ ID NO:5) to search through the ProDom database. Numerous matches against protein domains described as "channel protein calcium entry capacitative ionic transmembrane ion transport transient" and the like were identified.

#### Tissue Distribution of TLCC-4 and TLCC-5 mRNA

This example describes the tissue distribution of TLCC-4 mRNA, as determined by RT-PCR, and the tissue distribution of TLCC-4 and TLCC-5 mRNA as may be determined by Northern blot analysis.

Various cDNA libraries were analyzed by RT-PCR using a human TLCC-4-specific probe. From this analysis it was determined that TLCC-4 mRNA was expressed predominantly in the hypothalamus and skin. TLCC-4 mRNA was found in moderate levels in adipose and teste, and in low levels in skeletal muscle and brain (see Figure 3).

Northern blot hybridizations with the various RNA samples is performed under standard conditions and washed under stringent conditions, i.e., 0.2XSSC at 65°C. The DNA probe is radioactively labeled with 32P-dCTP (using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier). Filters containing human tissue mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto, CA) are probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

*In situ* hybridization experiments were performed using a human TLCC-4-specific probe indicating TLCC-4 expression in monkey brain (cortex, thalamus, caudate, and hippocampus), spinal cord, DRG and SRG neurons, and in hair follicles. *In*

*stii* hybridization with rat pain models indicated that TLCC-4 mRNA was down-regulated after chronic constriction injury, which causes persistent, spontaneous firing of neurons and results in pain. TLCC-4 mRNA was also down-regulated after treatment with clofibrate acid, a selective muscle toxin which produces muscle pain and inflammation.

#### EXAMPLE 2: EXPRESSION OF RECOMBINANT TLCC-4 AND TLCC-5 POLYPEPTIDES IN BACTERIAL CELLS

In this example, human TLCC-4 or human TLCC-5 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, TLCC-4 or TLCC-5 is fused to GST and this fusion polypeptide is expressed in *E. coli*, e.g., strain PEB199. Expression of the GST-TLCC-4 or GST-TLCC-5 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

#### EXAMPLE 3: EXPRESSION OF RECOMBINANT IC54420 POLYPEPTIDE IN COS CELLS

To express the human TLCC-4 or human TLCC-5 gene in COS cells, the pCDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire TLCC-4 or TLCC-5 polypeptide and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant polypeptide under the control of the CMV promoter.

To construct the plasmid, the human TLCC-4 or human TLCC-5 DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by approximately twenty nucleotides of the TLCC-4 or TLCC-5 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the TLCC-4 or TLCC-5 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are

digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the TLCC-4 or TLCC-5 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5 $\alpha$ , SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the human TLCC-4, pCDNA/Amp plasmid DNA or human TLCC-5, pCDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the IC54420 polypeptide is detected by radiolabelling (<sup>35</sup>S-methionine or <sup>35</sup>S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labelled for 8 hours with <sup>35</sup>S-methionine (or <sup>35</sup>S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the human TLCC-4 or human TLCC-5 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the TLCC-4 or TLCC-5 polypeptide is detected by radiolabelling and immunoprecipitation using a TLCC-4-specific or TLCC-5-specific monoclonal antibody.

#### EXAMPLE 4: TISSUE DISTRIBUTION OF HUMAN TLCC-4 mRNA USING TAQMAN™ ANALYSIS

This example describes the tissue distribution of human TLCC-4 mRNA in a variety of cells and tissues, as determined using the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA. The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest, *e.g.*, various human tissue samples, and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

As indicated in Figure 3, strong expression of TLCC-4 was detected in human brain (hypothalamus) and skin tissues. In addition, TLCC-4 expression was detected at moderate levels in adipose and testis tissues, and at low levels in the fetal heart, skeletal muscle, brain, and colon tissues. Pain human panel phase I and MP Phase 1.3.3 libraries

were also analyzed and it was determined that TLCC-4 was expressed at high levels in the brain, cortex, and testis, at moderate levels in the spinal cord, dorsal root ganglion (DRG), and the hypothalamus, and at low levels in the skin, placenta, small intestine, ovary, prostate epithelial cells, liver, skin (decubitus), colon tumor cells, and breast tumor cells (see Figures 12 and 13). Monkey libraries were also analyzed indicating that TLCC-4 was expressed at high levels in the monkey cortex and hairy skin, and at low levels in the monkey spinal cord (see Figures 14 and 15). Metabolic libraries were also analyzed demonstrating that TLCC-4 was expressed at high levels in adipose and brain tissues, and at low levels in differentiated adipocytes and pre-adipocytes, as well as in the hypothalamus, colon, small intestine, skeletal muscle, and liver tissues (see Figure 16).

#### EXAMPLE 5: REGULATION OF CALCIUM INFLUX THROUGH TLCC-4

This experiment describes the regulation of calcium influx through TLCC-4 in 911 cells as determined by Fluorometric Imaging Plate Reader experiments (FLIPR) (Molecular Devices Corp., Sunnyvale, CA).

The FLIPR is a screening tool for cell-based fluorescent assays which allows the simultaneous stimulation and measurement of separate cell populations in a high throughput format. Therefore, using this system, it is possible to quantify transient signals, such as the release of intracellular calcium, from cell populations, in parallel and in real time. The FLIPR contains chambers in which to hold the test plate and plates containing antagonists or agonists to be added to the test plate. The FLIPR utilizes an argon laser that provides discrete spectral lines spaced from approximately 350 to 530 nm. For use with fluorescent  $\text{Ca}^{2+}$  dyes, the 88-nm line of the laser is employed. The laser simultaneously illuminates the wells in a test plate. The image of each well in the plate is captured by a cooled charge coupled device (CCD) camera, which updates images once per second, if required, for the measurement of rapid calcium responses. Because both excitation and emission are read via the bottom of the plate, black-walled, transparent bottomed 96-well plates are used. Data captured by the CCD camera is converted to digital data and then transferred to a computer.

Briefly, a calcium indicator (*e.g.*, fluo-3/AM or Calcium Green-1/AM) was transferred to the culture medium. Because the FLIPR collects fluorescence from the bottom of the well, suspension cells require centrifugation to the base of the well following dye loading. Viable 911 cells were resuspended in loading medium and incubated for one hour. The cells were then centrifuged and resuspended with wash buffer. The cell suspension containing the dye was then aliquoted into each well of the black-walled, transparent bottomed 96-well plate and the plate was centrifuged. The

FLIPR assay was then carried out and the results analyzed. (If adherent cells are used, they may be plated at an appropriate density in the 96-well plates and cultured overnight. Dye may then be loaded and incubated).

Results show a constitutive calcium influx through TLCC-4 in 911 cells that were incubated with NMDG/0  $\text{Ca}^{2+}$  and stimulated afterwards with 5 mM  $\text{Ca}^{2+}$ .

#### Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed:

1. An isolated nucleic acid molecule selected from the group consisting of:
  - (a) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:1; and
  - (b) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:3; and
  - (c) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:4; and
  - (d) a nucleic acid molecule comprising the nucleotide sequence set forth in SEQ ID NO:6.
2. An isolated nucleic acid molecule selected from the group which encodes:
  - (a) a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2; and
  - (b) a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:5.
3. Isolated nucleic acid molecules comprising the nucleotide sequences contained in the plasmids deposited with ATCC® as Accession Numbers \_\_\_\_.
4. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2 or 5.
5. An isolated nucleic acid molecule selected from the group consisting of:
  - a) a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or a complement thereof;
  - b) a nucleic acid molecule comprising a fragment of at least 30 nucleotides of a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6, or a complement thereof;

c) a nucleic acid molecule which encodes a polypeptide comprising an amino acid sequence at least about 60% identical to the amino acid sequence of SEQ ID NO:2 or 5; and

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, wherein the fragment comprises at least 10 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2 or 5.

6. An isolated nucleic acid molecule which hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 under stringent conditions.

7. An isolated nucleic acid molecule comprising a nucleotide sequence which is complementary to the nucleotide sequence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.

8. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5, and a nucleotide sequence encoding a heterologous polypeptide.

9. A vector comprising the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5.

10. The vector of claim 9, which is an expression vector.

11. A host cell transfected with the expression vector of claim 10.

12. A method of producing a polypeptide comprising culturing the host cell of claim 11 in an appropriate culture medium to, thereby, produce the polypeptide.

13. An isolated polypeptide selected from the group consisting of:

a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, wherein the fragment comprises at least 10 contiguous amino acids of SEQ ID NO:2 or 5;

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2 or 5, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a complement of a nucleic acid molecule consisting of SEQ ID NO:1, 3, 4, or 6 under stringent conditions;

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 60% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, or 6; and

d) a polypeptide comprising an amino acid sequence which is at least 60% identical to the amino acid sequence of SEQ ID NO:2 or 5.

14. The isolated polypeptide of claim 13 comprising the amino acid sequence of SEQ ID NO:2 or 5.

15. The polypeptide of claim 13, further comprising heterologous amino acid sequences.

16. An antibody which selectively binds to a polypeptide of claim 13.

17. A method for detecting the presence of a polypeptide of claim 13 in a sample comprising:

a) contacting the sample with a compound which selectively binds to the polypeptide; and

b) determining whether the compound binds to the polypeptide in the sample to thereby detect the presence of a polypeptide of claim 13 in the sample.

18. The method of claim 17, wherein the compound which binds to the polypeptide is an antibody.

19. A kit comprising a compound which selectively binds to a polypeptide of claim 13 and instructions for use.

20. A method for detecting the presence of a nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in a sample comprising:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to a complement of the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to the complement of the nucleic acid molecule in the sample to thereby detect the presence of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 in the sample.

21. The method of claim 20, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.



22. A kit comprising a compound which selectively hybridizes to a complement of the nucleic acid molecule of any one of claims 1, 2, 3, 4, or 5 and instructions for use.

23. A method for identifying a compound which binds to a polypeptide of claim 13 comprising:

- a) contacting the polypeptide, or a cell expressing the polypeptide with a test compound; and
- b) determining whether the polypeptide binds to the test compound.

24. The method of claim 23, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

- a) detection of binding by direct detection of test compound/polypeptide binding;
- b) detection of binding using a competition binding assay; and
- c) detection of binding using an assay for TLCC-4 or TLCC-5 activity.

25. A method for modulating the activity of a polypeptide of claim 13 comprising contacting the polypeptide or a cell expressing the polypeptide with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

26. A method for identifying a compound which modulates the activity of a polypeptide of claim 13 comprising:

- a) contacting a polypeptide of claim 13 with a test compound; and
- b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

27. The method of claim 26, wherein said activity is modulation of nociception.

28. A method for identifying a compound which modulates pain comprising:

- a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and
- b) identifying the compound as a modulator of pain by determining the effect of the test compound on the activity of the polypeptide.

29. A method for identifying a compound which modulates nociception comprising:

- a) contacting the polypeptide of claim 13, or a cell expressing the polypeptide with a test compound; and
- b) identifying the compound as a modulator of nociception by determining the effect of the test compound on the activity of the polypeptide.

30. A method for treating a subject having a pain disorder comprising administering to the subject a TLCC-4 or TLCC-5 modulator, thereby treating said subject having a pain disorder.

31. A method for treating a subject having a pain disorder comprising administering to the subject a TLCC-4 or TLCC-5 modulator, wherein the TLCC-4 or TLCC-5 modulator is the modulator identified by the method of claim 26, thereby treating said subject having a pain disorder.

32. The method of claim 30, wherein the TLCC-4 or TLCC-5 modulator is a small molecule.

33. The method of claim 30, wherein said TLCC-4 or TLCC-5 modulator is administered in a pharmaceutically acceptable formulation.

34. The method of claim 30, wherein said TLCC-4 or TLCC-5 modulator is administered using a gene therapy vector.

35. The method of 30, wherein the TLCC-4 or TLCC-5 modulator is capable of modulating TLCC-4 or TLCC-5 polypeptide activity.

CCACGCGTCCGATTCACGCCATCCCTCTGCCTGCAATGAGAGCTTCCGCCGCCCTCAGCC  
ACAGTCCCACGCGGCGGTGGGCCCCAGACATCGGTGATCTCAGAGGAAAGGGTGTCC  
ACGACACGACGACGCTCAGACGCCATGAAGACCCACCCAGGAGATGGTGCGCTTCAT  
GGCAGAGAGATTTGCGGCCCACTGGGACCTCGCTCTCCAGAGAGAGCGCCGCG  
GGAGATACCCCAACAGAGAGTGCACATCTTCTTGAGATGAGAAGGTTGTAAC  
CAAGCCACAGTTGGCAAGACCTGCTCTGTTCTCCAGGCCATGGATTCCCAAT  
CCGCGAGTGCACTCTGTGAATCTGTATGACATGGAATGGACTCCCCAGTCTCTCAAGATG  
TTGGACAGAGACCCCATCCCAACAGCCACGAGTGGCACTGCGCCACAGGAGGACGA  
GAGAGGAAAGAGCGGCGCTGGAAGGCCATCTTGACGCGTCTGTCGAGGCTGCGT  
GGAGGAGTTGTGAGATTTGGTGTGGAGCTGCGAGAGCTTTGACGGCGGCCATGATGA  
GATGATGGCTTCTCTCATGACACAGCTGACGGCTCTCGACGGCGGCCATGCTCT  
CTTTCTGAGAGACGACATCTCGGCGAGTTCATCACCGCAGGTACACAGAGAGGC  
CTATGAGGGCAGCGGCTGAACATGCGCATGACGGCGGAGGGGACATCCACGCC  
CTCTCTATCCGCGCGCGCGAGCTCAAGCGGACGCCAAGGGCGCTTCTTCAGCC  
CAAGTACCAACAGAAAGTCTTACTTTCGTGTGAGACGCGCTCGCTGGCAGCATGCCAC  
CAACGACCGGAGATTGCGACGCTGTATGAGACGACGACGAGGACATCATCTCGG  
GGACTCAGAGGCAACATCTCTTCAGCGCTGACGCGCTGCGCGGAGAGATCTCAAGC  
GAGAAATGACTTTGTGAAGGCAATGACATCATCTTACTGCGAGTGGCAAGTGGGA  
BCTGAGAGACTCGCAACACGAATGCCCTACCGCGCTCAGCTGGCGGCAAGATGGG  
CAAGCGGAGATCTGTGAATGATCATCTCAATGCTGAGATCAAGAGAAAGCGGCTCGGAG  
CCTGTCCAGGAATTCACCGACGAGGCTGACGAGCGGTGATCTTCTTGAAGCGTCTGG  
CACCAAGTGGACAGGACACGACGACATCGAGTGTGGGATATCACTCTCTACACCA  
CATGCAACCGGCATGAGATGCTGACCTTGAGCGCCCTGCACGCTCGCATGATGAA  
GTGAGAAAGTTTGGCAACGACATTTCTTTCTGCTGTGTTTATTTCTCTACAA  
CATCACTTGAACCTCTGCTACTACCTCGCGCGCGGAGGAGGCCATCCGCAACC  
CTTGGCCTGACGCAAGATGCGGTGGCTGAGCTCTAGGAGATGTTTGTGCTCAT  
CTGGGCCATGCTATCTGTGAAGGGCATATGCCATCTTTCTTGAAGCGTCTGG  
TCTGACGATACTCTCTCGAATGCGGTCTCCACTGTGTTTATTCGAAGCTGTCT  
TGTGATCTGTCTGTCTGTACTTGTTCCTTACAAAGACTCTCGCTGCTGCTGT  
GTGCGCAATGGCGCTGGCGGCAATCTGTACTACTACGCGGGTGTCCAGTGCAT  
GGGATGTGACGCTGATGATCAGAGAGTCAATTTGAGATGATTTGTGAAGTCTGT  
TGATATATCGTGTTTTGTGTGATTTGAGTAGCTTCGCGCTCGTCAATCGAAGTG  
TCCCAACAGCAAGGACTGACGCTCTAGGCACTTCTGAGCGCAATGCTGAGACT  
CTTCAAGTCAACATAGCGCTGGGTGATGCAACATCAGACGAACTCAAGTATGCCAT  
TCTCTTTGTTCTGCTCATCACTATGTCTACTCTCACTTGTCTCTCTCTCAACAT  
CGCCTTCGACAGAGCGGACCATCTTGGAGTTTGAGAAATGTTACCAAGACGCATCTG  
GAGCAATCCGATTTGGAGAGCTGTCAAAGTGGCGGAGAGATGTTCCGACTGTGTT  
GCGATCAATAGGTGATGAGTCAATGAAGACGACGCTCTCTCTCTTACGAGAA  
CCGGCGGCTGTAAAGACGACAGATTTCAACAAATCCAAATCTTCCAGGACAAACGA  
CAAAACCACTCTCAATGCAATTTGAAGAGTGTGAGAAATTCGCGGAGTGTAGAA  
GAGAACCCAGTCTGTGTGCTGCGCTGTCTGCTGCGCTGCAGCGGAGTCCGCTG  
CTGTGACAGAGGCTTTGAGGATGATGAGATGCTGCTGCGCTGCAGCGGAGTCCGAC  
CAGCTCTGCTGCTGACTCATGGGTGTGTGAACATGAGGCGAGTGTCAATTTGTCTG  
ATGGGAAACATCTGTGATTTGTGTACTTGTCAACAGCACTGTGTAACTACGACGAC  
AGCATCTGAGCGCTGGAGCGCTCTGAGATCTCCCGGTGAAGCTCTGTTTACCAATT  
CGAGTTCGCTTGGCTGGGAGATGATGATGGCGGGAAAGGGGACAGCTGTGAGGAGCA  
GGAGAGGAGTCTTTCTCTCCGACTGTCTCCGCTGCTCCGCAACCCGACCCACAGAT  
TGTAACATCTTCTGCTGTGCTGGTTCGCTGATTTGTGGAGACGCGTGAATCCCAT  
CCGAGATGTGCTGGGAGCGAGGATGAGCTGTGATATCTGGGAGACGCGTGAATCCCAT  
CTGGAAACACCTTCGGGACCTTTGGGGAAGAAGGAGACATCTTGGAGTAGAGCTC  
CTGACACCCCAAGCTTCAACTCTGACGCTGAGAGATGTTTGTAGTACGATAATTA  
CAGAGGGTGTAACTGCTCAATACGGAACAGACTGCGGCTGTGAGAAATGAAATTA  
ATGAGTTTGTGATTTTAAAGTGGAGACAGACTGGGCGCTTGGGGCTGTATGTAAAGTA

## FIGURE 1

MKAHPKBNVPLMGURVLAASGNPAVLPEKKKPAEITPTTKSAHFLEIEGEPNPPTVAKTS  
 PPVFSMDNISIRQCTISGNCMDQSPQODOVTFPTSPNPSPSQAQAKESRCKKRRLL  
 KRIFAAVSGCVLLELVELLVQLCRRHDEVDVPLMKHLTASTQKTLCKMALKLNTIN  
 PNTKIVIRLILAPABENDILGRFNAEYETAYEQGLIAALITERRQDIAALIIAAGAD  
 NAWAKGAFNPVKQHBGFQETPLAALACTNQPEIVOLLMEHQOTITSROSRCNNLIT  
 HALVYVAEDPKTDQFVKYMDHMLLSGNGWELETITRRNDOGLTPQLAAKSAELIKYI  
 LSREIMKSLASLSKRFOTWAYGVSSSLYITDNYDTVITVNTINIDNRHML  
 TLEPLHTLMLMKKKAKMFLSCFPPFVNTILVSYRPREZEAIPHPLATKNG  
 WLLQJLQRMFVLWANCISKEGAIILFLRPSQMGITFQVAVLVFLVFLV  
 LFAYKSYLACLVLMALGANNMLYTFQSGMQGYSVM1QKV1LHDVLKLVFYVIFLGL  
 FOVALASLEKCPKDKCCSSVGSFSDAVLEPKLITGLGDLNIQNSKYPILFLPLIT  
 YVILTVPLLMMLLALQEDVTENVSNSKSEIRIHLQARATITLEFQHLFWLRSRFRMGE  
 CKVAEDFALCLRNENKVTX

## FIGURE 1

Input file Pbh53920c.seq; Output file Pbh53920c.lta  
Sequence length 3042

A V G G C R W T E D V E P A E V R E K H 20  
ATG GTT GGA GCA TCC AGG TGG ACA GAC GTG GAG CCT GCA GAA GGA MAG GAA ATG 60  
S P R A A R L S M R M R N D T L D S T 40  
TCC TTT CGG GCA ACC AGC CTC ACC ATG AGG AAC AGA ACC ACT CTG GAC ACC T 120  
R T L Y S S A S R S T D L S Y S E S D L 60  
CGG ACC CTG TAC TCC AGC GCG TCT CGG AGC ACA CTG TCT TAC AGT GAA GAC TTG 180  
V N P I O A N P R R E C V F P T K D S 80  
CTG AAT TTT CAA GCA AAT TTT MAG AAA GAA TGT GTG TTC TTT ACC AAA CAT TCC 240  
K A T E N V C K G G Y A C G S O H M E G T 100  
AAG GGC ACC GAG AAT GTG ACC TCC AGC TCC AGC ACC GAC ACC GAC ATG GAA GGC ACC 300  
O I N O S E K W N Y K K H T R E F P T D 320  
CAG ATC ATC GAG AAT TCC AAC TCC AAC TCC AAC TCC AAC TCC AAC TCC AAC TCC 360  
A P O D I O P E T L G K K G K Y I R L S 140  
GCC TTT CGG GAT ATT CAG TTT GAG ACA CTG GCG MAG AAC TCC GAC TAT ATA GGT CTG TCC 420  
C D T D A E I L Y E L L T O H W H L K T 160  
TCC GAC ACC GCA GCA ATC CTT TAC TAC GAG CTG ACC CAG TCC TCC CAG CTG AAA ACA 480  
P N L V I S V T G G A K N F A L K P R H 180  
CCC AAC CTG ATT TCT GTG ACC GCG GCC ACC MAG AAC TTC GCC CTG AAC CCG GCG ATG 540  
R K I P S R L I Y I A O S K G A T W AT T C K 200  
GCC MAG ATC TTC AGC CGG CTC ATC ATC GCG CAG TCC AAA GGT GGT TCC AT TCC 600  
O O T H Y G C L N K Y L G E V V R D N T I 220  
GAA GGC ACC CAT TAT GGC CTG ATG MAG TAC CTC CGG GAG GTG GTG ACA GAT AAC ACC ATC 660  
S R S S E E N I V A I O I A A W G M V S 240  
AGC AGG ACT TCA GAG GAG AAT ATT GTG GGC ATT GTG GGC ATG GCA GCT TCG GGC ATG TCC 720  
N R D T L I R N C D A E G Y F L A O Y L 260  
AAC CGG GAC CTC ATC AGG AAT TCC GAT GCT GAG GGC TAT TTT TTA GCC CAG TAC GTT 780  
H D D P T R D P L Y I L D N N H T H L L 280  
ATG GAT GAC TTC ACA GAA GAT DCA CTG TAT ATC CTG GAC AAC AAC ACA CAT TTG CTG 840  
L V D N G C H G H P T V E A K L R N O L 300  
CTC GTG GAC AAT GGC TGT CAT GCA GCA CCG ACT GTG GAA GCA CTC CGG AAT CAG CTG 960  
E K Y I S E R T I O D S N Y G G K I P I 320  
GAG MAG TAT TCC GAG GGC ACT ATT CAA GAT TCC AAC TAT GGT GCG MAG ATC CCC ATT 960  
V C F A O G O G K E T L K A I N T S I K 340  
GTG TCT TTT GCC CAA GGA GGT GGA AAA GAG ACT TTG AAA GGC ATC AAC TCC ATC AAA 1020  
H R I P C V V V E O S G O I A D V I A S 360  
AAT AAA ATT CCT TGT GTG GTG GAG GGC CTG GAG GGC CAG CTG GAT GAT GTG ATC GGT ACC 1080  
L V E V E D A L T S S A V R E K L V R Y 380  
CTG GTG GAG GAG GAT GGC GTG ACA TCT TCT GGC GTG MAG MAG MAG MAG MAG MAG MAG 1140  
L P R T V S R L P E E E T E S M I K M L 400  
TTA CCC CGC CGC CTC CGG CTC CGG GAG GAG ACT GAG AGT TCC AAC AAA TCC CTC 1200  
R E I L E G S H L L T V I K H E E A G D 420  
AAA GAA AAT CTC GAA TCT TCT CAG CTA TTA ACA GTT ATT AAA ATG GAA GAA GGT GCG GAT 1260

Figure 2

3/23

E I V S N A I S Y A L Y K A F S T S E Q 440  
GAA ATT GTG AGC AAT GGC ATC TCC TAC CTA TAC AAA GGC TTC ACC ACC AGT GAG CAA 1320  
D K D N M N O Q L K L L E W N O L D 460  
GAC MAG GAT AAC TGG AAT GCG CAG CTG MAG CTT CTG CTG TCG AAC CAG CTG GAC TTA 1380  
A N D E I P T N D R R W E K S K P R L R 480  
GCC AAT GAT AAT TTT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT 1440  
D T I I O V T W L E N G R I K V E S K D 500  
GAC ACA ATA ATC CAG CTC ACA TGG CTG GAA AAT GGT AGA ATC AAG GTT GAG AGC AAA GAT 1500  
V T D G R A S S H M L V V L K S A D L Q 520  
GTG ACT GAC GGC AAA GGC TCT TCT CAT ATG CTG GAT GAT GAT GAT GAT GAT GAT GAT GAT 540  
E V H P T A L I K D R P K F V R L P L E 560  
GAA GTG ATG TTT AGC GCT CTC ATA MAG GAC AGA CCC MAG TTT GTG CCG CTC TTT CTG GAG 1620  
N G L N L R K F L T H D V L T E L F S N 580  
AAT GGC TTG AAC CTA CGG MAG TTT CTC ACC CAT GAT GTG CTC ACT GAA CTC TTC TCC AAC 1680  
H F S T L V Y R N L O I A K N S Y N D A 580  
CAC TTC AGC ACC CTT CTG CAG AAT GGC ATC GGC MAG AAT TCC TAT TAT AAT GAT GGC 1740  
L L T P V W K L V A N F R G F R K E D 600  
CTC CTC AGC TTT GTG TGG AAA CTG GTT GCG AAC TTC CGA ACA GGC TTC CGG MAG GAA GAC 1800  
R N G R D E H D I E L H D V S P I T R H 620  
AGA AAT GGC CGG GAG GAG ATC GAT ATA GAA CTC GAC GAT TCC CTT ACT CCG CAC 1860  
P L O A L F I W A I L O N K E L S K V 640  
CCC CTC CAA GCT CTC ATC ATG TGG GAT ATT CTT CAG AAT MAG MAG GAA CTC TCC AAA CTC 1920  
I W E O T R G C T L A A L G A S K L L K 660  
ATT TGG GAG CAG ACC AGC AGC TCC ACT CTC GCA GGC CTG GCA GGC ACC MAG CTT CTG MAG 1980  
T L A K V R N D I N A A G E S E L A N 680  
ACT CTG GGC AAA GTG MAG AAC GAT AAT GCT GCG GAG GAG TCC GAG GAG CTG CTG AAT 2040  
E Y L T R A V G E S T V W N A V V G A D 700  
GAG TAC CTG ACC CGG GCT GTT GGT GAG TCC ACA GTG TGG AAT GCT GTG GAG GGC GAT 2100  
L P C G T D I A S G T H R P D G G E L F 720  
CTG CCA TGT TAC ACA GAC ATT GGC AGC ACT AAC ACA GAT GAT GAT GAT GAT GAT GAT 2160  
T E C Y S S D E D L A E O L L V Y S C E 740  
ACT GAG TGT TAC AGC AGC GAT GAA GAC TTG GCA GAA CAG CTG CTG GAT TAT TCC TCT GAA 2220  
A W G G S N C L E L A V E A T D O H F I 760  
GCT TGG GGT GGA AGC AAC TGT CTG GAG CTG GCG GTG GAG GGC ACA GAC CAG CAT TTC ATC 2280  
A O P G V O N F L S K O W Y G E I S R D 780  
GCC CAG CCT GCG CTC CAG AAT TTT CTT TCT MAG CAA TGG TAT GGA GAG AAT TCC CCA GAC 2340  
T K N W K I I L C L F I I P L V G C G F 800  
AAC MAG MAG TGG MAG AAT ATC CTG TGT CTG TTT ATT ATA CCC TTG GTG GGC TGT GGT TTT 800  
V S P R K K P V D R K R K L L W Y V A 820  
TCA TCA TTT AGG MAG AAA CCA GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT GAT 840  
P F T S P P V V V F S W N V V V F Y I A F L 860  
TTC TTC ACC TCC CCC TTC GTG GTG TTC TCC TGG AAT GTG GTG TTC TAC AAC GGC TTC CTC 2520  
L L F A Y V L L M D F H S V P P P E L 860

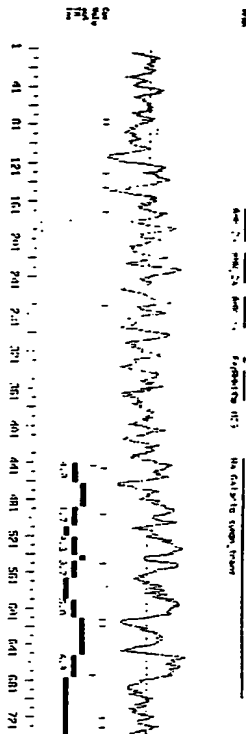
Figure 2

4/23

Figure 3

CTG CTG TTT GCC TAC GTG CTG CTG ATG GAT TTC CAT TCG GTG CCA CAC CCC CCC GAG CTG 2580  
V L Y S L V F V L F C D E V R Q G R P A 880  
GTC CTG TAC TCG CTG GTC TTT GTC CTG TGT GAT GAA GTG AGA CAG GGC CCG CCT 2640  
A P S A Q P A K P T P T R N S I W P A S 900  
GCT CCC AGT GCG GGG CCC GGC AGC CCC ACC CCG AAC TCC ATC TCG CCC GCA AGC 2700  
S T R S P G S R S R H S F H T S L Q A E 920  
TCC ACA CCG AGC CCC GGT TCC CCG TCA CCG CAC TTC TTC CAC ACT TCC CTG CAA GCT GAG 2760  
G A S S G L G Q P R K G L P Q C S G G L 940  
GCT CCC AGC TCT GGC CAG CCC AGA AAG GGG CTC CCA CAG TCC AGC GGT GGG CTG 2820  
K G S S S A A K V G A Q A E E V P R A S 960  
AAG GGC TCC TCA AGT GCC GCC AAA GTG GGA GCC CAG GCA GAG GTG CCG AGA GCA AGC 2880  
E G C E D C Q H A V T S Q K R K T A M D 980  
GAG GGC TGT GAG GAC TCC CAG CAC GCT GTC ACC TCT CAG AAG CCT AAG ACA GCA ATG GAC 2940  
Q T D E D L F P Y G A F Y Q F L M I S R 1000  
CAA ACA GAC GAA GAT CTC TTC CCC TAT GGA GCA TTC TAC CAG TTC CTG ATG ATT TCC AGG 3000  
S F R G E E M S I G K Q H \* 1014  
AGO TTT CGA GGA GAG GAG ATG AGC ATC GGC AAG CAG CAC TAA 3042

Figure 2



>48000FL  
 MKAPREKRVNPLAKGVAAAPSGNPAVUPKPKPAETTPKXSAFLEIEGEPNPVAVTS  
 PVPEKPMDSNIRQICSGNCDMDSPQSDQDVETETSPNSPSPADAKEORRRRLK  
 KRTFAVSEGVSEVELVELLOELCRHDEDPDPLAKHLTAEDGKCLAKLALININ  
 PRTKAIYRITLLAFABENDILORFTNAETBEAYEGQALNIAIERQGDIALIAGAD  
 VNAHAKAPFNPKYOHGEGFYGETPLALACTNQPEIYVOLLMEHQDITSDSGNNIL  
 HALVTVAEDPKTONDFVKNVYDMLLRSGNWELETTNNDGLTQLAAMGKALIKYI  
 LSREIKERLRSLSRKFTDNAVGVSSSLYDVTNVDITDINSVLEITVNTNIDRHEML  
 TLEPLHTLHMKKKFAKMFSLSCFYFYVITLTVSYRPREEAIHPPLALATHMG  
 WQLLGMVPLIWMACISVKEGIALFLABSDIQLSDAMPHPVFPQIQAIVITISVPLV  
 LFAIKETLACTVLAMALGMANMLYTRGQSGMGMVIOKVLHNDVAKLFVYIVFLG  
 PGVALALIEKCPKNDKCSYGSFSDAVLEFLTLTGLDLNIQONSXKPLFLFLIT  
 VYIITFVLNMLIALMGSTVENSKESERIVRLQBARITLFEKMLPEVLSRRNGEL  
 CVAEDDFRLCLAINKVKMTX

FIGURE 4

Protein Family / Domain Matches, HMMer version 2  
 Searching for complete domains in PFAM  
 hmmpfam - search a single seq against HMM database  
 HMMER 2.1.1 (Dec 1998)  
 Copyright (C) 1992-1998 Washington University School of Medicine  
 HMMER is freely distributed under the GNU General Public License (GPL).  
 HMM file: /prod/dm/seganal/pfam/pfam.2/pfam  
 Sequence file: /prod/dm/wspace/orfana/oa-script.28060.seq  
 Query: 48000FL

Scores for sequence family classification (score includes all domains):			
Model	Description	Score	E-
ank	ank repeat	79.0	1e-
19	4		
ion_trans	ion transport protein	34.5	2.5e-
06	1		
Peptidase_M29	Thermophilic metalloprotease (M29)	0.9	
4	1		
Na_Galacto_symp	Sodium:galactoside symporter family	-153.9	
8.4	1		

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
ank	1/4	167	202	1	33	1.6	1.4e+02
ank	2/4	214	246	1	33	30.6	3.5e-05
ank	3/4	261	294	1	33	27.9	0.00024
Peptidase_M29	1/1	325	331	402	408	0.9	4
ank	4/4	340	372	1	33	18.6	0.15
ion_trans	1/1	510	677	1	223	34.5	2.5e-06
Na_Galacto_symp	1/1	435	699	1	285	-153.9	8.4

Alignments of top-scoring domains:

ank: domain 1 of 4, from 167 to 202: score 1.6, E = 1.4e+02  
 \*\*->dgtrpLhlaarng...hlevkllleagadvarck<-+  
 G T L + A n + + + e + v + + L L + + +  
 48000FL 167 TGTCLAKALININPTKEIVRILAFABENDILGR 202  
 ank: domain 2 of 4, from 214 to 246: score 30.6, E = 3.5e-05  
 \*\*->dgtrpLhlaarng...hlevkllleagadvarck<-+  
 +G T L + A + + + + + + + L L + aGAdvna K  
 48000FL 214 EGGTALNIAIERROGDIALIAGADVNAHAK 246  
 ank: domain 3 of 4, from 261 to 294: score 27.9, E = 0.00024  
 \*\*->dgtrpLhlaarng...hlevkllleagadvarck<-+  
 G T P L I A A + + + e + v + + L L + + + d + r d  
 48000FL 261 FGRTPLAACTNQPEIYVOLLMEHQDITSDRS 294

FIGURES









Fb52920c hTRP7 MWGCRWTEBEPAPAVEKEMGFRAALSNR-NRNDYLSNLTYSAGRSNDLSYBSBD  
 MEPSALRAGGEBOEGEGELPRRYTDLGMVSNLR-RSNSSLSFSMQLCPFGNDRKOS-  
 \* \* \* \* \*  
 Fb52920c hTRP7 LWNFIQANFKRECEVFTKDSKATEN---VKCCYASQNHMEGT---QINOSKNNYKKA  
 LSSWJPENIKKKECYFVESRSLDAGKVNOCGTHQHEBEPHPTFOGTDPKKA  
 \* \* \* \* \*  
 Fb52920c hTRP7 TKEPPTDAFGDIQFETLCK-GKTRLSCTDPAELVELLJOMHLKTNLIVISYTGAK  
 VOEMPTDAFGDIQFETLCK-GKTRLSCTDPAELVELLJOMHLKTNLIVISYTGAK  
 \* \* \* \* \*  
 Fb52920c hTRP7 NFALKPRNKIFSR-LIYIAGSKAMITGTHYGMKYGEEVVRONTISRSSEEN-IVA  
 NFEMKPRLSKISIRRCGLVAVQTTGAMITGSHGKNGKQNGEAVRDFSSSYKEBELT  
 \* \* \* \* \*  
 Fb52920c hTRP7 IOLANQWNSNDTLIRNODAGVFLADYLDPTDPLYLIDNNHTLLLVNDCGHAP  
 IGVATWGTVHRREGLIH---PTGSFPABYIIDEQOCNLTCLDSNHSFILLVNDGTHGY  
 \* \* \* \* \*  
 Fb52920c hTRP7 TVEAKLRNOLERYISERTIODSNYGSKIPIVCAAGGCKETLKAINTSINQKIPCVVVG  
 GVEITLKTREKPISEQTERGVAIKIPIVGVVLEGGPGTLITIDMAITNGPCVVG  
 \* \* \* \* \*  
 Fb52920c hTRP7 SGQIDVIAIV--EVEDALTSAAVEKLVNLPRTVSLPEETESIMKLEIECSH  
 SGVADVIAQVANLPVSQ-ITISLOOKLSVFGQMFETFSERIVEMTAKIODIYARRQ  
 \* \* \* \* \*  
 Fb52920c hTRP7 LLTIVIKNEAGDEIVSNATSVALLYNA-PTSEBODKONNGOLKLLBWNQDLANDBEIT  
 LLTIVREGKDGODVVALIQAALNARSQDHFGEHMDHQLKLVAMNVDLAREEIPM  
 \* \* \* \* \*  
 Fb52920c hTRP7 NDRRWIKS--KEPLADTIIO-----VTWLENGRIKVESQDVDTGAKSSMVLVLSADL  
 DEMQKPSDLPTMTALISNKBFWKJPLENG---VOLKEPVMTDILLYLYXENLDESCL  
 \* \* \* \* \*  
 Fb52920c hTRP7 OEVMFT-ALIKDRPKFVRLPLENGMLKRPGLTHDVTLSFSNHPTSLVYNNQIATKSNY  
 FHEKQKVLVED-PE-RPACAPAPRLQMHVADVRELLG-DFTQPLYP--RRRHRDR-  
 \* \* \* \* \*  
 Fb52920c hTRP7 DALLTTPWKLIVNFRGPRKEDNGENDIDELHNVSPITRPIQALEPIMAILONKEELS  
 LRLLLPVPHVKLAVO-GVS--LYRBSGH-VT-PTMDPIRDLIIVAYONREELA  
 \* \* \* \* \*  
 Fb52920c hTRP7 KVIWETRGCTLAALASQLKTLAKVNDINAGSEBELANETLTRAVGESTVMANAVG  
 GIIMWSQDCTIAALACSKILKELSKESBDTDSSEBELALAEYEHRAIG-----V-  
 \* \* \* \* \*  
 Fb52920c hTRP7 ADLPCGTDIASGTHRPGGELFTECYSDDEDLAEQLLVYRCEANGGSCLELAVEATDQH  
 -----FTECYRDEBRQKLLTVSEAMGKTYCLOLALKAQMK  
 \* \* \* \* \*  
 Fb52920c hTRP7 FIAOPGVQNFILSKWYGEISRDTKQNMKIYCLPTIPLVGCGVSRKRPVDIGKKLAMY  
 FVSHGGLQFLIVKMGQGLSDVNDQKREVLICLALPPLILNLSRBERLDVGTTPARA  
 \* \* \* \* \*  
 Fb52920c hTRP7 VAFTS#VVVSNVNVVYIAVILPAIVILDPHVSFHPPELVLSYPLACDVRQGR  
 RAFTARQDUZHNILISRTAFCLFAIYVAMDQPVVSWCBCHILYMLSEICEROLF  
 \* \* \* \* \*

FIGURE 10

## multiple sequence alignment

```

--JUSVAL W (1.74)
Pbh52920c
hMELASTATIN
MYIRSVYDTKPDLSLLHLMVKDQLELPKLLSVHGGLOQNFEMQPKLKQVFGKGLIKAAMT
--
Pbh52920c
hMELASTATIN
TGAMIFTGGUSTGVISHVGDAKDHSSKSGRCAIGIAPMGIVENKEDLVGKVOTRYVC
--
Pbh52920c
hMELASTATIN
KQSPRAARLGNR-----RRNDTLDST-----RTLYSSAS-RGTDLSYSR-SDLV
TMSNPLSKLSVLNNSHTFILDADNGTIGKYGAELRLLEKHISLQKINTRELQGGVPLV
--
Pbh52920c
hMELASTATIN
NFIQA----NFKKRECVFTKDSKATENVCK-CGYAQS-----QHM---EGTQINOSE-KWN
GLVVEGGPNVVISIVLEYLQEGPPIPVVICDGSGRASDILSFAHKYCEGGIINESLREQI
--
Pbh52920c
hMELASTATIN
YKHTKEFPPTDARFDIQ-FETLG---KKGKYI---RLSCDITDAEILYELLQHWH---LK
LVTTQKTFNTNKAQSHQLFAIIMECKMKKELVTVMGSEGOQDIENAILTALLKGTNVS
--
Pbh52920c
hMELASTATIN
TPN-LVIVSVTGG-----AKNFALKPRMRKIPFRLYTAQSKGAW-----ILT-GGTH
APDQLSLALAWNRVDIARSQIFVFGP-HMTPLGSLAPPTDSKATEKPKPPMATTKGRC
--
Pbh52920c
hMELASTATIN
YGLMKYLGEV---VRONTISRSEE-NIV-AIGIAANG-MVSNR-D---TLIRN-CDAG
KGGKKGKGYKEVEEETDPRKIETLLMWNALQFQMLDALVLDVDFVKLIENGVMQNH
--
Pbh52920c
hMELASTATIN
YFLAQLMDDF-TRD-P-----LYILDN---NHT---HLLVDNGCHGHPTVEAKLRNL
FLTIPLREELYNTRGAPPNTLHLVDRDKKSNLPDPYHISLIDIGLVLEYLMGAVRCNY
--
Pbh52920c
hMELASTATIN
EKYISERTIQDSNYGKGIPIVCPAOG3-----GKE-----TLKAINTSIKN-----
TRKN-PRTLNMLFGPKPKALKLLGHWDEDPAPKGGKKKKKKKEEIDIDVDDPAVSRF
--
Pbh52920c
hMELASTATIN
KIP--CVWVEGSGQIADVIASLVEV--EDALTSSAVKEKLVRELP-----TVSRL
QYPPHELMVWAVLMKQKNAFLWQGESMAKALVACKLYKANAHESSESDLVDDISQD
--
Pbh52920c
hMELASTATIN
PEEETESWIKWLKEILECS--HLLTVIKMEERAGD-EIVSNAISVAL-----YKAFS--TS
LDNNSKDFQQLALELDQSYKHDEQIAMKLLTYELKWNSTCLKLAFAAKHRDFIAHTC
--
Pbh52920c
hMELASTATIN
EQ--DKDNMNGQLK-----LLEWNOLD-----LAN-D-----E
SQMLLTDMNGRLMRKPNGLKVIINGILLPPTILFLEPRTYDFSYQTSKENEDGKEKE
--
Pbh52920c
hMELASTATIN
IFT--N-D---RRW-EKSKRLRDTI---IQVTLENG-RIKVESKQVT-DGKASSMLV
ENTDNADAGSRKGEENEHKQKRSIPIGTKICEFYNAPIVKNFNFYITISYLGILLNFYV
--
Pbh52920c
hMELASTATIN
VLKASD-----LQEWNTALIKORP-KFVRLFLEN--G-IN--LRKFLTHDV-LTELPFN-
ILVRMDGWPLOEMIVISIVSLALEKIBELIMSEPKLSQKIKVMLOEYWNITDLVAIS
--
Pbh52920c
hMELASTATIN
HPST-LVYRMLQIAK-NSYNDALL-----TFWKLKLVNFRGFKED-----RNGDEMDIE
TFMIGAILR-LQNOPTMGYGRVIYVDIIFNTVRLDIFGVNKKILGYVVMIGKWMIDML

```

Figure 11

```

LHDVS-----PITR-----HP-----LQALF-----IWAIF-----Q-----N
YFVIMLVLMVFGVARQAILHPEEKPSKMLARNIFMYPWMIYGEVPADQIDLYAMEIN
--
K-----KE-----LSKVI---W---EQTRGC-----TLAALQASKLLKLTAKVQNDIN
PFCGNLDEEGKLPFCIPGAWLTPALMACTLLVANILLNLLTAVPNTFFEFKISIN
--
AAGESR--ELANEYLTR-----AVGESTVNAVVGADLPCGTDIAGTTHRPD-QGELPT
QVWKFORQLLIMTFHDPVLPPEMILSHIYIINRLSGCRKKREGQOEQRDLGLKFL
--
-----ECYSS-----DEDLA-----EQLLVYCEANGGNSCLELAVEATDQHF
SDEELKLHEFEQVQEHFREKEDEQSSDERIRVTSERVENMSMLEEINER--ETV
--
IAQPGVQNP-LSKQWYGBIS-RDT---KNWKILCLFIPLVG-----CG-FVSFRKKPV
MKYS-LQITVDLURLAQLEELSNRWVALENLAGIDRDLIOAKSRASSECEATYLLRSSI
--
--D-----KH-----KLLWYVYVFTPTSPFV-VFSNNVVFYIA---FLLLFAYVLLMDF
NSADGYSLYR YHFNGBELLFEDTSLSTSGTGVRKKKTSFRIKEEKDVKTHLVPEQNSL
--
H-----SVP---H-----PPELVLYSLVFLFCDEVROGR-----PAAPSA---
HLSLGTSATPDGSHLAVDDDLKNAEESKLPDQIGISKDDBRQTDSKKEETISPSLNKT
--
--OPAK-----PTPTRN---SI-NPASS---TR-SPGS-----RSRUSPHTS
DVTHGQDSVDQNTQLTVEETNIEGTISYPLETKITRFPDETINACKTWKSRSFYISR
--
LQ--AEGA-----SSGLQOPRKGLPQCS--GGLKGSSS-----AAKYGAQAEVPRASE
GRKLVGVGNQDVEYSSITDQQLTTEMQCVQKITRSHSTDIPYIVSEAAVQAQEKQFAD
--
GCEDCQNAVTSQREK-----TAMDQT-DEDLFPYG--AFYQP-LMISRPGBE---MSI
-MQDEHHVAEAIPIRIPRSLTITDRNGENLLSVKPDQTLGPPSLRSKSLHGHPRNYSI
--
-GKQ-----H-----
QKLDKRGSHASSVSSLVIVSGMTAEKKYKKEKASTETEC

```

Figure 11

48000 Panel Human Phase I

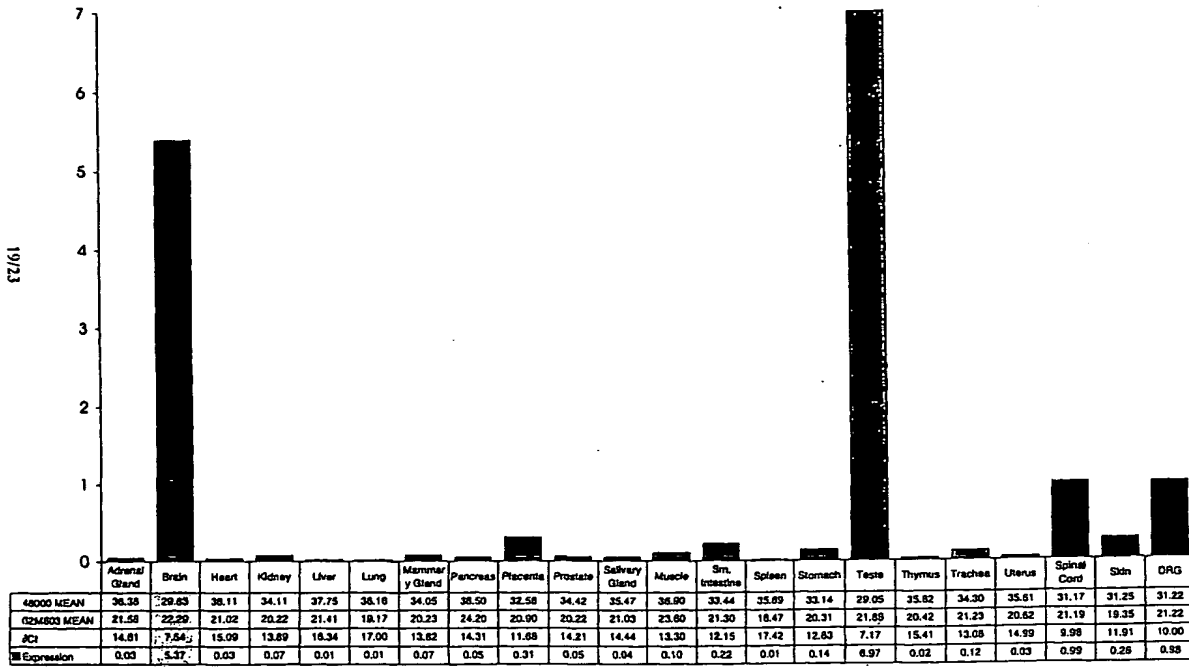


Figure 12

Phase 1.3.3 Expression of 48000.1 w/02

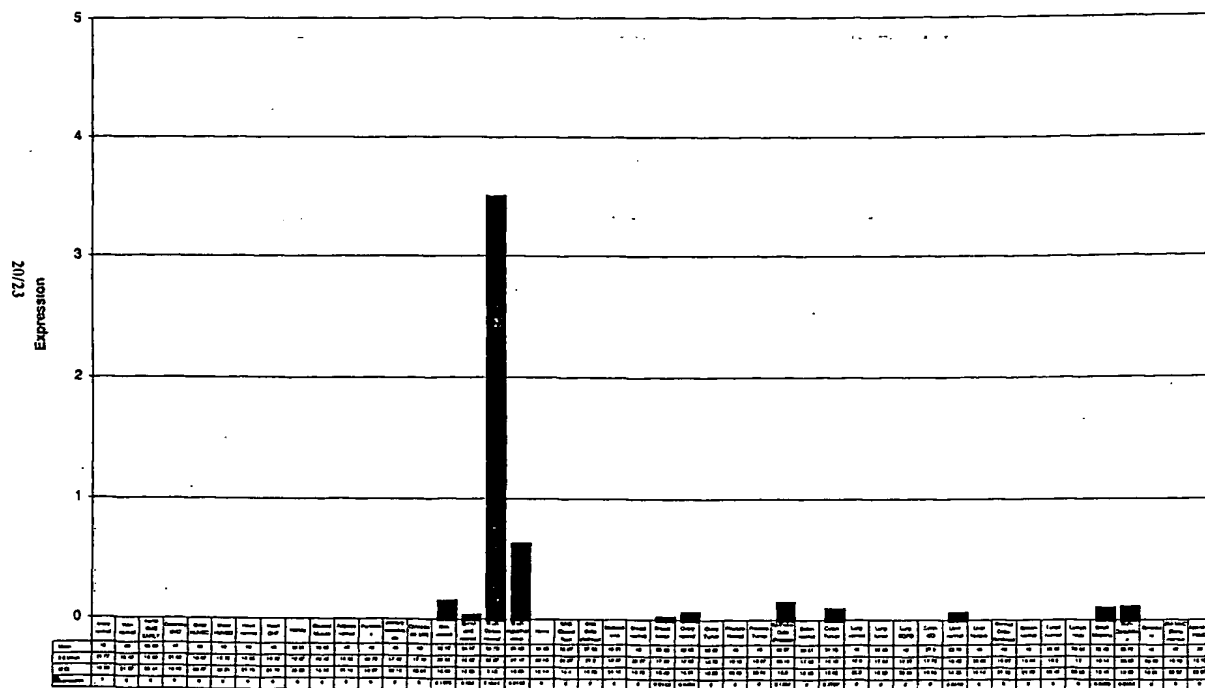


FIGURE 13

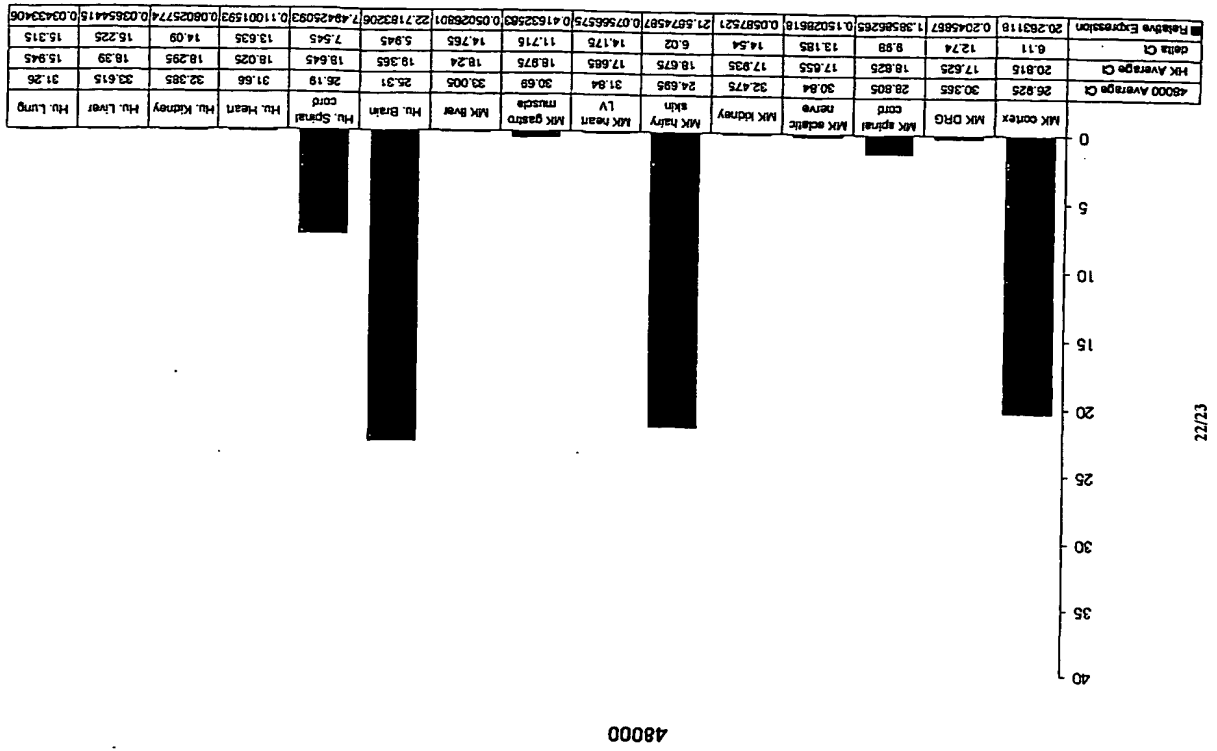


FIGURE 15

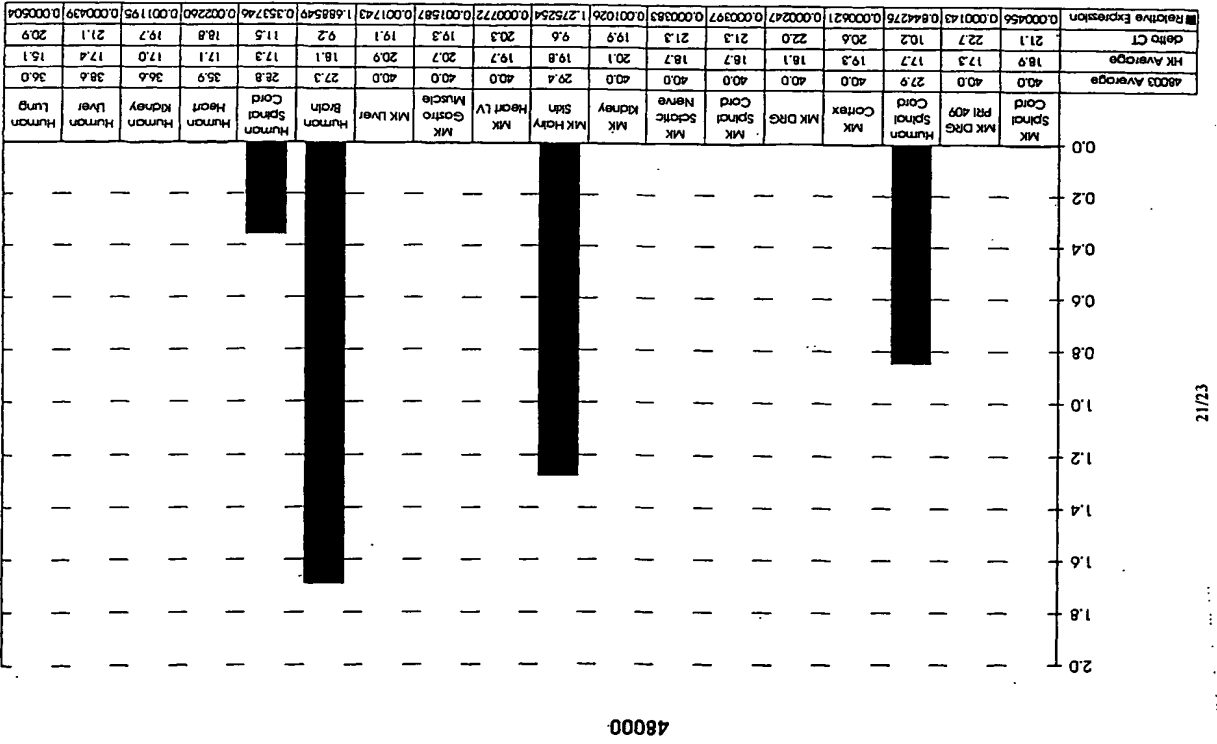


FIGURE 14

## 48000 Metabolic II

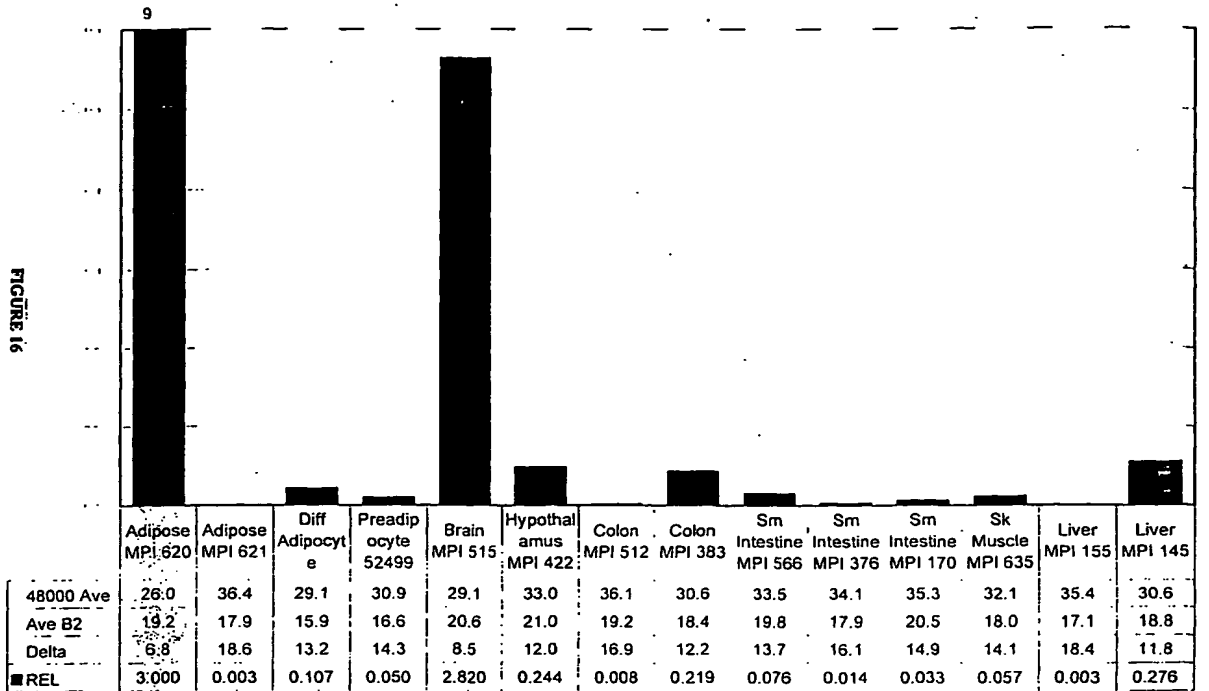


FIGURE 16

23/23

## SEQUENCE LISTING

<110> Millennium Pharmaceuticals, Inc., et al.  
 <120> 48000 AND 52920, NOVEL HUMAN CALCIUM CHANNELS AND USES THEREOF  
 <130> NMI-170FC  
 <150> US 60/214,176  
 <151> 2000-06-26  
 <160> 6  
 <170> FastSeq for Windows Version 4.0  
 <210> 1  
 <211> 4586  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> CDS  
 <222> (146) ... (2368)  
 <400> 1  
 ccacgctcc gattccagcc atccctctgc ctgcgaatgag agcttccgc cgctccagcc 60  
 acagctccac ccggggcgtc tggccacag acatcggtg atccagggc aaggttgc 120  
 acgacaccc agactctac cagcc atg aag gcc cac ccc aag gag atg gtg 172  
 Met Lys Ala His Pro Lys Glu Met Val 1 5  
 cct ctc atg ggc aag aga gtt gct gcc ccc agt ggg aac cct gcc gtc 220  
 Pro Leu Met Gly Lys Arg Val Ala Ala Pro Ser Gly Asn Pro Ala Val 10 25  
 ctg cca gag aag agg ccg gcg gag atc acc ccc aca aag aag agt gca 268  
 Leu Pro Glu Lys Arg Pro Ala Glu Ile Thr Pro Thr Lys Lys Ser Ala 30 40  
 cac ttc ctg gag ata gaa ggg ttt gaa ccc aac ccc aca gtt gcc 316  
 His Phe Phe Leu Glu Ile Glu Gly Phe Glu Pro Asn Pro Thr Val Ala 45 55  
 aag acc tct cct cct gtc ttc tcc aag ccc atg gat tcc aac atc cgg 364  
 Lys Thr Ser Pro Pro Val Phe Ser Lys Pro Met Asp Ser Asn Ile Arg 60 70  
 cag tgc atc tct ggt aac tgt gat gac atg gac tcc ccc cag tct cct 412  
 Gln Cys Ile Ser Gly Asn Cys Asp Asp Met Asp Ser Pro Gln Ser Pro 75 85  
 caa gat gat gtg aca gag acc cca tcc aat ccc aac agc ccc agt gca 460  
 Gln Asp Asp Val Thr Glu Thr Pro Ser Asn Pro Asn Ser Pro Ser Ala 90 105  
 cag ctg gcc aag gaa gag cag agg aag aag aag agg cgg ctg aag aag 508  
 Gln Leu Ala Lys Glu Glu Gln Arg Arg Lys Lys Arg Arg Leu Lys Lys 110 120  
 cgc atc ttt gca gcc gtg tct gag ggc tgc gtg gag gag ttg gta gag 556

Arg Ile Phe Ala Ala Val Ser Glu Gly Cys Val Glu Glu Leu Val Glu  
125 130 135  
604  
ttg ctg gta gta ctg gag cag ctt tgc agg cgg cgc cat gat gag gat  
Leu Leu Val Glu Leu Glu Leu Cys Arg Arg Arg His Asp Glu Asp  
140 145 150  
652  
gtg cct gag ttc ctc atg cac aag ctg acg gcc tcc gag acg ggg aag  
Val Pro Asp Phe Leu Met His Lys Leu Thr Ala Ser Asp Thr Gly Lys  
155 160 165  
700  
acc tgc ctg atg aag gcc ttg tta aac atc aac ccc aac acc aag gag  
Thr Cys Leu Met Lys Ala Leu Leu Asn Ile Asn Pro Asn Thr Lys Glu  
170 175 180 185  
748  
atc gtg cgg atc ctg ctt gcc ttt gct gaa gag aac gat atc ctg gcc  
Ile Val Arg Ile Leu Leu Ala Phe Ala Glu Glu Asn Asp Ile Leu Gly  
190 195 200  
796  
agg ttc atc aac gcc gag tac aca gag gag gcc tat gaa ggg cag acg  
Arg Phe Ile Asn Ala Glu Tyr Thr Glu Glu Ala Tyr Glu Gly Gln Thr  
205 210 215  
844  
gag ctg aac atc gcc atc gag cgg cgg cag ggg gag atc gca gcc ctg  
Ala Leu Asn Ile Ala Ile Glu Arg Arg Gln Gly Asp Ile Ala Ala Leu  
220 225 230  
892  
ctc atc gcc gcc ggc gcc gac gtc aac gcg cac gcc aag ggg gcc ttc  
Leu Ile Ala Ala Gly Ala Asp Val Asn Ala His Ala Lys Gly Ala Phe  
235 240 245  
940  
ttc aac ccc aag tac caa cac gaa ggc ttc tac ttc ggt gag acg ccc  
Phe Asn Pro Lys Tyr Glu His Glu Gly Phe Tyr Phe Gly Glu Thr Pro  
250 255 260 265  
988  
ctg gcc ctg gca gca tgc acc aac cag ccc gag att gtg cag ctg ctg  
Leu Ala Leu Ala Ala Cys Thr Asn Gln Pro Glu Ile Val Gln Leu Leu  
270 275 280  
atg gag cac gag cag acg gag atc acc tgc cgg gag tca cga ggc aac  
1036  
Met Glu His Glu Gln Thr Asp Ile Thr Ser Arg Asp Ser Arg Gly Asn  
285 290 295  
aac atc ctt cac gcc ctg gtg acc gtg gcc gag ttc aag acg cag  
1084  
Asn Ile Leu His Ala Leu Val Thr Val Ala Glu Asp Phe Lys Thr Gln  
300 305 310  
aat gag ttt gtg aag cgc atg tac gag atg atc cta ctg cgg agt ggc  
1132  
Asn Asp Phe Val Lys Arg Met Tyr Asp Met Ile Leu Leu Arg Ser Gly  
315 320 325  
aac tgg gag ctg gag acc act cgc aac aac gat ggc ctc acg cgg ctg  
1180  
Asn Trp Glu Leu Glu Thr Thr Arg Asn Asn Asp Gly Leu Thr Pro Leu  
330 335 340 345  
cag ctg gcc gcc aag atg ggc aag gcg gag atc ctg aag tac atc ctc  
1228

Gln Leu Ala Ala Lys Met Gly Lys Ala Glu Ile Leu Lys Tyr Ile Leu  
350 355 360  
agt cgt gag atc aag gag aag cgg ctc cgg agc ctg tcc agg aag ttc  
1276  
Ser Arg Glu Ile Lys Glu Lys Arg Leu Arg Ser Leu Ser Arg Lys Phe  
365 370 375  
acc gag tgg gcg tac gga ccc gtg tca tcc tcc ctc tac gag ctc acc  
1324  
Thr Asp Trp Ala Tyr Gly Pro Val Ser Ser Ser Leu Tyr Asp Leu Thr  
380 385 390  
aac gtg gag acc acc acg gag aac tca gtg ctg gaa atc act gtc tac  
1372  
Asn Val Asp Thr Thr Thr Asp Asn Ser Val Leu Glu Ile Thr Val Tyr  
395 400 405  
aac acc aac atc gag aac cgg cat gag atg ctg acc ctg gag cgg ctg  
1420  
Asn Thr Asn Ile Asp Asn Arg His Glu Met Leu Thr Leu Glu Pro Leu  
410 415 420 425  
cac acg ctg ctg cat atg aag tgg aag aag ttt gcc aag cac atg ttc  
1488  
His Thr Leu Leu His Met Lys Trp Lys Lys Phe Ala Lys His Met Phe  
430 435 440  
ttt ctg tcc ttc tgc ttt tat ttc ttc tac aac atc acc ctg acc ctc  
1516  
Phe Leu Ser Phe Cys Phe Tyr Phe Phe Tyr Asn Ile Thr Leu Thr Leu  
445 450 455  
gtc tgc tac tac cgc ccc cgg gag gag gcc atc cgg cac ccc ttg  
1564  
Val Ser Tyr Tyr Arg Pro Arg Glu Glu Ala Ile Pro His Pro Leu  
460 465 470  
gcc ctg acg cac aag atg ggg tgg ctg cag ctc cta ggg agg atg ttt  
1612  
Ala Leu Thr His Lys Lys Met Gly Trp Leu Gln Leu Leu Gly Arg Met Phe  
475 480 485  
gtg ctc atc tgg gcc atg tgc atc tct gtg aaa gag ggc att gcc atc  
1660  
Val Leu Ile Trp Ala Met Cys Ile Ser Val Lys Glu Gly Ile Ala Ile  
490 495 500 505  
ttc ctg ctg aga ccc tgc gat ctg cag tcc atc ctc tgc gat gcc tgg  
1708  
Phe Leu Leu Arg Pro Ser Asp Leu Gln Ser Ile Leu Ser Asp Ala Trp  
510 515 520  
ttc cac ttt gtc ttt ttt atc caa gct gtg ctt gtg ata ctg tct gtc  
1756  
Phe His Phe Val Phe Phe Ile Gln Ala Val Leu Val Ile Leu Ser Val  
525 530 535  
ttc ttg tac ttg ttt gcc tac aaa gag tac ctc gcc tgc ctc gtg ctg  
1804

4

Phe Leu Tyr Leu Phe Ala Tyr Lys Glu Tyr Leu Ala Cys Leu Val Leu  
 540 545 550  
 gcc atg gcc ctg gcc tgg gcc aac atg ctc tac tat acg cgg ggt ttc  
 1852  
 Ala Met Ala Leu Gly Trp Ala Asn Met Leu Tyr Tyr Thr Arg Gly Phe  
 555 560 565  
 cag tcc atg gcc atg tac agc gtc atg atc cag aag gtc att ttg cat  
 1900  
 Glu Ser Met Gly Met Tyr Ser Val Met Ile Glu Lys Val Ile Leu His  
 570 575 580 585  
 gat gtc ctg aag ttc ttg ttt gta tat atc gtc ttt ttg ctt gga ttt  
 1948  
 Asp Val Leu Lys Phe Leu Phe Val Tyr Ile Val Phe Leu Gly Phe  
 590 595 600  
 gga gta gcc ttg gcc tgg ctg atc gag aag tgt ccc aaa gac aac aag  
 1996  
 Gly Val Ala Leu Ala Ser Leu Ile Glu Lys Cys Pro Lys Asp Asn Lys  
 605 610 615  
 gac tgc agc tcc tac gcc agc ttc agt gac gca gty ctg gaa ctc ttc  
 2044  
 Asp Cys Ser Ser Tyr Gly Ser Phe Ser Asp Ala Val Leu Glu Leu Phe  
 620 625 630  
 aag ctg aac ata gcc ctg ggt gat ctg aac atc cag cag aac tcc aag  
 2092  
 Lys Leu Thr Ile Gly Leu Gly Asp Leu Asn Ile Glu Glu Asn Ser Lys  
 635 640 645  
 tat ccc att ctc ttt ctg ttc ctg ctc atc acc tat gtc atc ctc acc  
 2140  
 Tyr Pro Ile Leu Phe Leu Phe Leu Leu Ile Thr Tyr Val Ile Leu Thr  
 650 655 660 665  
 ttt gtc ctc ctc ctc aac atg ctc att gct ctg atg gcc gag act gty  
 2188  
 Phe Val Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val  
 670 675 680  
 gag aac gtc tcc aag gag agc gaa cgc atc tgg cgc ctg cag aga gcc  
 2236  
 Glu Asn Val Ser Lys Glu Ser Glu Arg Ile Trp Arg Leu Glu Ala  
 685 690 695  
 aag acc atc ttg gag ttc gag aaa atg. tta cca gaa tgg ctg aag agc  
 2284  
 Arg Thr Ile Leu Glu Phe Glu Lys Met Leu Pro Glu Trp Leu Arg Ser  
 700 705 710  
 aga ttc cgg atg gga gag ctg tgc aaa gty gcc gag gat gat ttc cga  
 2332  
 Arg Phe Arg Met Gly Glu Leu Cys Lys Val Ala Glu Asp Asp Phe Arg  
 715 720 725 730  
 ctg tgt ttg cgg atc aat gag gty aag tgg act gaa tgaagacgc  
 2378

5

Leu Cys Leu Arg Ile Asn Glu Val Lys Trp Thr Glu  
 730 735 740  
 acgtctcctt ccttaacgaa gaaccgggc ctgtaagc aacgatctt aacaaatcc  
 2438  
 aagattcttc caggaaacac agcaaaacca ctcccaatgc atttgaaga gtcgaagaat  
 2498  
 tcccgaaac ctccgtgttag aagcgaaac cagaatcgtt gtcgcgtgc gctgtcgtc  
 2558  
 gctgcagcgc gattcaccca ctctgtcag aagagcttgc aaggaatg gattcgcgt  
 2618  
 ctctgtgcct agaagcagag tgcacctgc tgcctcaatgc tcatgtgtg tctgaactga  
 2678  
 gggtcagttg tcaattgtc tgaatggaa acatcctgga ttltgttact tggcaaacg  
 2738  
 ctgtgttaaa cctaacgcca gcagcagctt ggaagctggg agctctccta agtcccggt  
 2798  
 gaagcctctg gtttaacca ttgcagtcg gcttgcctgg gagaatgga tggcgggaaa  
 2858  
 gggtcagcag tcttgagag caggagagag agtcttccct cctgcacagc tcccgcgtca  
 2918  
 gcccaaccc cagcccaaac attgtaccat ctctctgct gtcactggtt tgcctgaatt  
 2978  
 tgtggagag ccgtgatccc atcccaagat gtcgcgggga cggaggtga ctgatattc  
 3038  
 tggggagaga ggggaatgcg ctctgaaac acccttcagg aacccttcgg ggaagaagag  
 3098  
 accatccttg gattgaacgt cccctgacac cccaagcttc aaactgtctc aagctgaag  
 3158  
 atgttttag tagcataatt aacacaggtt tttaacttgc aatacgaaa agaatltca  
 3218  
 gttggaatg aaaaactact caatgaagtt tgtgaattta aaagtggaga cagactggg  
 3278  
 gctttgggc tgaatgtaag tattatara ttgacctcag ggtgcacgaa gaagacaaa  
 3338  
 aagctttctt tcaacacac aaaaagctgc atgagacat ccgggcaagt cctgtcggc  
 3398  
 cgcgcgatac tgggtgaaag gtccctgctc ttltcctgt cctgacctca cagtacgca  
 3458  
 tgcctgtg ctggatcgt ggccttcgt gaagcagaaa tagcagctgc tcatcgata  
 3518  
 tcaacttga actcagcagt tagtcgcata cctcagtcag tctcagtygg gaaatltac  
 3578  
 aaaaagctc aactgcttgc gtaagaatga ttutttt aatttlaact gtaattttg  
 3638  
 aagctgaag ggaagctgtt gagaagaaag catttgcca gaatttgagc ttatttttag  
 3698  
 gtccctgctc tctgatgttc tctttctgaa atgacagga gtcagctcgg gggcagag  
 3758  
 tgaagtggag acggaagat ttccaagty actggggccg aaaccacag aaaaactcact  
 3818  
 ctgcgcgctt tatctgtgaa aagatltac gtaaaaatgt tcaagtygaa atabaanaa  
 3878  
 tagtaacct aatgttlaat cttaaatgc agaataagaa atttgcctt cagataacat  
 3938  
 ggagataag aatltacat ggccttgagc aaactgaaga gtcggggcct aagattgcac  
 3998  
 tctgggcag ttctctgac ctgtgtgcca gactagctgg acagatagag  
 4058  
 acttctgac cctgatggg ccgattggg agaagtgggc tgggtgtgc aggttcaaa  
 4118

atccacagca gcccctgccc tcccagctga cccaggaggt aatcgcgtgc tctaagccac  
 4178  
 adtggtcggg gctgggcatg ggcctctgga gaagagaaga ttggaggaga actgtcctag  
 4238  
 aggcaggagg agcagatgtg ttccagaatg ggcagaatta ggaaattgag aaagattttg  
 4298  
 gctcaacaga atccagcaac tgcctcagat gtggagatg tttaagcaga agctgggtga  
 4358  
 gaactaatg aggaatgttg ttgaaatagg tcaattgaag aagtttaagg tcccttttag  
 4418  
 ctggagatt gtacaatca gcatccaca tctggagtta gctaccgca ttaagcctga  
 4478  
 acagacatct tggctgaaa gaaagtgtt tggattcatg atgccaagct ccacactatg  
 4538  
 gagctgggaa ttccagaatt gctttgactc agatattaat ggagaag  
 4586

&lt;210&gt; 2

&lt;211&gt; 741

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 2

Met Lys Ala His Pro Lys Glu Met Val Pro Leu Met Gly Lys Arg Val  
 1 5 10 15  
 Ala Ala Pro Ser Gly Asn Pro Ala Val Leu Pro Glu Lys Arg Pro Ala  
 20 25 30  
 Glu Ile Thr Pro Thr Lys Lys Ser Ala His Phe Phe Leu Glu Ile Glu  
 35 40 45  
 Gly Phe Glu Pro Asn Pro Thr Val Ala Lys Thr Ser Pro Pro Val Phe  
 50 55 60  
 Ser Lys Pro Met Asp Ser Asn Ile Arg Gln Cys Ile Ser Gly Asn Cys  
 65 70 75  
 Asp Asp Met Asp Ser Pro Gln Ser Pro Gln Asp Asp Val Thr Glu Thr  
 80 85 90 95  
 Pro Ser Asn Pro Asn Ser Pro Ser Ala Gln Leu Ala Lys Glu Glu Gln  
 100 105 110  
 Arg Arg Lys Lys Arg Arg Leu Lys Lys Arg Ile Phe Ala Ala Val Ser  
 115 120 125  
 Glu Gly Cys Val Glu Glu Leu Val Glu Leu Leu Val Glu Leu Gln Glu  
 130 135 140  
 Leu Cys Arg Arg Arg His Asp Glu Asp Val Pro Asp Phe Leu Met His  
 145 150 155  
 Lys Leu Thr Ala Ser Asp Thr Gly Lys Thr Cys Leu Met Lys Ala Leu  
 160 165 170 175  
 Leu Asn Ile Asn Pro Asn Thr Lys Glu Ile Val Arg Ile Leu Leu Ala  
 180 185 190  
 Phe Ala Glu Glu Asn Asp Ile Leu Gly Arg Phe Ile Asn Ala Glu Tyr  
 195 200 205  
 Thr Glu Glu Ala Tyr Glu Gly Gln Thr Ala Leu Asn Ile Ala Ile Glu  
 210 215 220  
 Arg Arg Gln Gly Asp Ile Ala Ala Leu Leu Ile Ala Ala Gly Ala Asp  
 225 230 235 240  
 Val Asn Ala His Ala Lys Gly Ala Phe Phe Asn Pro Lys Tyr Gln His  
 245 250 255  
 Glu Gly Phe Tyr Phe Gly Glu Thr Pro Leu Ala Leu Ala Cys Thr  
 260 265 270  
 Asn Gln Pro Glu Ile Val Gln Leu Leu Met Glu His Glu Gln Thr Asp  
 275 280 285  
 Ile Thr Ser Arg Asp Ser Arg Gly Asn Ile Leu His Ala Leu Val  
 290 295 300

Thr Val Ala Glu Asp Phe Lys Thr Gln Asn Asp Phe Val Lys Arg Met  
 305 310 315  
 Tyr Asp Met Ile Leu Leu Arg Ser Gly Asn Trp Glu Leu Glu Thr Thr  
 320 325 330 335  
 Arg Asn Asn Asp Gly Leu Thr Pro Leu Gln Leu Ala Ala Lys Met Gly  
 340 345 350  
 Lys Ala Glu Ile Leu Lys Tyr Ile Leu Ser Arg Glu Ile Lys Glu Lys  
 355 360 365  
 Arg Leu Arg Ser Leu Ser Arg Lys Phe Thr Asp Trp Ala Tyr Gly Pro  
 370 375 380  
 Val Ser Ser Ser Leu Tyr Asp Leu Thr Asn Val Asp Thr Thr Thr Asp  
 385 390 395 400  
 Asn Ser Val Leu Glu Ile Thr Val Tyr Asn Thr Asn Ile Asp Asn Arg  
 405 410 415  
 His Glu Met Leu Thr Leu Glu Pro Leu His Thr Leu Leu His Met Lys  
 420 425 430  
 Trp Lys Lys Phe Ala Lys His Met Phe Phe Leu Ser Phe Cys Phe Tyr  
 435 440 445  
 Phe Phe Tyr Asn Ile Thr Leu Thr Leu Val Ser Tyr Tyr Arg Pro Arg  
 450 455 460  
 Glu Glu Glu Ala Ile Pro His Pro Leu Ala Leu Thr His Lys Met Gly  
 465 470 475  
 Trp Leu Gln Leu Leu Gly Arg Met Phe Val Leu Ile Trp Ala Met Cys  
 480 485 490 495  
 Ile Ser Val Lys Glu Gly Ile Ala Ile Phe Leu Leu Arg Pro Ser Asp  
 500 505 510  
 Leu Gln Ser Ile Leu Ser Asp Ala Trp Phe His Phe Val Phe Ile  
 515 520 525  
 Gln Ala Val Leu Val Ile Leu Ser Val Phe Leu Tyr Leu Phe Ala Tyr  
 530 535 540  
 Lys Glu Tyr Leu Ala Cys Leu Val Leu Ala Met Ala Leu Gly Trp Ala  
 545 550 555  
 Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln Ser Met Gly Met Tyr Ser  
 560 565 570 575  
 Val Met Ile Gln Lys Val Ile Leu His Asp Val Leu Lys Phe Leu Phe  
 580 585 590  
 Val Tyr Ile Val Phe Leu Leu Gly Phe Gly Val Ala Leu Ala Ser Leu  
 595 600 605  
 Ile Glu Lys Cys Pro Lys Asp Asn Lys Asp Cys Ser Ser Tyr Gly Ser  
 610 615 620  
 Phe Ser Asp Ala Val Leu Glu Leu Phe Lys Leu Thr Ile Gly Leu Gly  
 625 630 635 640  
 Asp Leu Asn Ile Gln Gln Asn Ser Lys Tyr Pro Ile Leu Phe Leu Phe  
 645 650 655  
 Leu Leu Ile Thr Tyr Val Ile Leu Thr Phe Val Leu Leu Asn Met  
 660 665 670  
 Leu Ile Ala Leu Met Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser  
 675 680 685  
 Glu Arg Ile Trp Arg Leu Gln Arg Ala Arg Thr Ile Leu Glu Phe Glu  
 690 695 700  
 Lys Met Leu Pro Glu Trp Leu Arg Ser Arg Phe Arg Met Gly Glu Leu  
 705 710 715 720  
 Cys Lys Val Ala Glu Asp Asp Phe Arg Leu Cys Leu Arg Ile Asn Glu  
 725 730 735  
 Val Lys Trp Thr Glu  
 740

&lt;210&gt; 3

&lt;211&gt; 2223

&lt;212&gt; DNA



<213> Homo sapiens  
 <220>  
 <221> CDS  
 <222> (1)... (2223)  
 <400> 3  
 atg aag gcc cac ccc aag gag atg gtg cct ctc atg gcc aag aga gtt 48  
 Met Lys Ala His Pro Lys Glu Met Val Pro Leu Met Gly Lys Arg Val 15  
 1 5  
 gcc gcc ccc agt ggg aac cct gcc gtc ctg cca gag aag agg ccg gcg 96  
 Ala Ala Pro Ser Gly Asn Pro Ala Val Leu Pro Glu Lys Arg Pro Ala 30  
 20 25  
 gag atc acc ccc aca aag aag agt gca cac ttc ttc ctg gag ata gaa 144  
 Glu Ile Thr Pro Thr Lys Lys Ser Ala His Phe Phe Leu Glu Ile Glu 45  
 35 40  
 ggg ttg gaa ccc aac ccc aca gtt gcc aag acc tct cct cct gtc ttc 192  
 Gly Phe Glu Pro Asn Pro Thr Val Ala Lys Thr Ser Pro Pro Val Phe 60  
 50 55  
 tcc aag ccc atg gat tcc aac atc cag cag tgc atc tct ggt aac tgt 240  
 Ser Lys Pro Met Asp Ser Asn Ile Arg Gln Cys Ile Ser Gly Asn Cys 80  
 65 70  
 gat gac atg gac tcc ccc cag tct cct caa gat gat gtc aca gag acc 288  
 Asp Asp Met Asp Ser Pro Gln Ser Pro Gln Asp Asp Val Thr Gly Thr 95  
 85 90  
 cca tcc aat ccc aac agc ccc agt gca cag ctg gcc aag gaa gag cag 336  
 Pro Ser Asn Pro Asn Ser Pro Ser Ala Gln Leu Ala Lys Glu Glu Gln 105  
 100  
 aag agg aaa aag agg cgg ctg aag aag cgc atc ttt gca gcc gtg tct 384  
 Arg Arg Lys Lys Arg Arg Leu Lys Lys Arg Ile Phe Ala Ala Val Ser 125  
 115 120  
 gag ggc tgc gtg gag gag tgg gta gag ttg ctg gtg gag ctg cag gag 432  
 Glu Gly Cys Val Glu Glu Glu Val Leu Leu Val Glu Leu Gln Glu 140  
 130 135  
 ctt tgc aag cgg cgc cat gat gat gag gat gtg cct gac ttc ctc atg cac 480  
 Leu Cys Arg Arg Arg His Asp Glu Asp Val Pro Asp Phe Leu Met His 160  
 145 150  
 aag ctg acg gcc tcc gac acg gag aag acc tgc ctg atg aag gcc ttg 528  
 Lys Leu Thr Ala Ser Asp Thr Gly Lys Thr Cys Leu Met Lys Ala Leu 175  
 165 170  
 tta aac atc aac ccc aac acc aag gag atc gtg cgg atc ctg ctt gcc 576  
 Leu Asn Ile Asn Pro Asn Thr Lys Glu Ile Val Arg Ile Leu Leu Ala 190  
 180 185  
 ttt gct gaa gag aac gac atc ctg ggc agg ttc atc aac gcc gag tac 624  
 Phe Ala Glu Glu Asn Asp Ile Leu Gly Arg Phe Ile Asn Ala Glu Tyr 205  
 195 200

aca gag gag gcc tat gaa ggg cag acg gcc ctg aac atc gcc atc gag 672  
 Thr Glu Glu Ala Tyr Glu Glu Thr Ala Leu Asn Ile Ala Ile Glu 215  
 210  
 cgg cgg cag ggg gac atc gca gcc ctg ctc atc gcc gcc gcc gcc gac 720  
 Arg Arg Gln Gly Asp Ile Ala Ala Leu Leu Ile Ala Ala Gly Ala Asp 240  
 225 230  
 gtc aac gcg cac gcc aag ggg gcc ttc ttc aac ccc aag tac caa cac 768  
 Val Asn Ala His Ala Lys Gly Ala Phe Phe Asn Pro Lys Tyr Glu His 255  
 245 250  
 gaa gcc ttc tac ttc ggt gag acg ccc ctg gcc ctg gca gca tgc acc 816  
 Glu Gly Phe Tyr Phe Gly Glu Thr Pro Leu Ala Leu Ala Ala Cys Thr 270  
 260 265  
 aac cag ccc gag att gtg cag ctg ctg atg gag cac gag cag acg gac 864  
 Asn Gln Pro Glu Ile Val Gln Leu Leu Met Glu His Glu Gln Thr Asp 285  
 275 280  
 atc acc tcg cgg gac tca cga gcc aac aac atc ctt cac gcc ctg gtg 912  
 Ile Thr Ser Arg Asp Ser Arg Gly Asn Asn Ile Leu His Ala Leu Val 300  
 290 295  
 acc gtg gcc gag gac ttc aag acg cag aat gac ttt gtg aag cgc atg 960  
 Thr Val Ala Glu Asp Phe Lys Thr Gln Asn Asp Phe Val Lys Arg Met 315  
 305 310  
 tac gac atg atc cta ctg cgg agt ggc aac tgg gag ctg gag acc act 1008  
 Tyr Asp Met Ile Leu Leu Arg Ser Gly Asn Trp Glu Leu Glu Thr Thr 335  
 325 330  
 cgc aac aac gat ggc ctg acg ccg ctg cag ctg gcc gcc aag atg ggc 1056  
 Arg Asn Asn Asp Gly Leu Thr Pro Leu Gln Leu Ala Lys Met Gly 350  
 340 345  
 aag gcg gag atc ctg aag tac atc ctc agt cgt gag atc aag gag aag 1104  
 Lys Ala Glu Ile Leu Lys Tyr Ile Leu Ser Arg Glu Ile Lys Glu Lys 365  
 355 360  
 cgg ctc cgg agc ctg tcc agg aag ttc acc gac tgg gcg tac gaa ccc 1132  
 Arg Leu Arg Ser Leu Ser Arg Lys Phe Thr Asp Trp Ala Tyr Gly Pro 380  
 370 375  
 gtg tca tcc tcc ctc tac gac ctc acc aac gtg gag acc acc acg gac 1200  
 Val Ser Ser Ser Leu Tyr Asp Leu Thr Asn Val Asp Thr Thr Thr Asp 400  
 385 390  
 aac tca gtg ctg gaa atc act gtc tac aac acc aac atc gac aac ccg 1248  
 Asn Ser Val Leu Glu Ile Thr Val Tyr Asn Thr Asn Ile Asp Asn Arg 415  
 405 410  
 cat gag atg ctg acc ctc gag ccg ctg cac acg ctg cat atg aag 1296  
 1290

His Glu Met Leu Thr Leu Glu Pro Leu His Thr Leu Leu His Met Lys  
420 425

tgg aag aag ttt gcc aag cac atg ttc ttt ctg tcc ttc tgc ttt tat  
1344

Trp Lys Lys Phe Ala Lys His Met Phe Phe Leu Ser Phe Cys Phe Tyr  
435 440 445

ttc ttc tac aac atc acc ctg acc ctg tgc tgc tac tac cgc ccc cgg  
1392

Phe Phe Tyr Asn Ile Thr Leu Thr Leu Val Ser Tyr Tyr Arg Pro Arg  
450 455 460

gag gag gag gcc atc cag cac ccc ttg gcc ctg acg cac aag atg ggg  
1440

Glu Glu Glu Ala Ile Pro His Pro Leu Ala Leu Thr His Lys Met Gly  
465 470 475 480

tgg ctg cag ctg cta ggg agg atg ttt gtg ctg atc tgg gcc atg tgc  
1488

Trp Leu Gln Leu Leu Gly Arg Met Phe Val Leu Ile Trp Ala Met Cys  
485 490 495

atc tct gtg aaa gag gcc att gcc atc ttc ctg ctg aga ccc tcg gat  
1536

Ile Ser Val Lys Glu Gly Ile Ala Ile Phe Leu Leu Arg Pro Ser Asp  
500 505 510

ctg cag tcc atc ctg tgc gat gcc tgg ttc cac ttt gtc ttt ttt atc  
1584

Leu Gln Ser Ile Leu Ser Asp Ala Trp Phe His Phe Val Phe Phe Ile  
515 520 525

caa gct gtg ctt gtg ata ctg tct gtc ttc ttg tac ttg ttt gcc tac  
1632

Gln Ala Val Leu Val Ile Leu Ser Val Phe Leu Tyr Leu Phe Ala Tyr  
530 535 540

aaa gag tac ctg gcc tgc ctg ctg gcc atg gcc ctg gcc tgg gcg  
1680

Lys Glu Tyr Leu Ala Cys Leu Val Leu Ala Met Ala Leu Lys Trp Ala  
545 550 555 560

aac atg ctg tac tat acg cgg ggt ttc cag tcc atg gcc atg tac agc  
1728

Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln Ser Met Gly Met Tyr Ser  
565 570 575

gtc atg atc cag aag gtc att ttg cat gat gtt ctg aag ttc ttg ttt  
1776

Val Met Ile Gln Lys Val Ile Leu His Asp Val Leu Lys Phe Leu Phe  
580 585 590

gta tat atc gtg ttt ttg ctt gga ttt gga gta gcc ttg gcc tcg ctg  
1824

Val Tyr Ile Val Phe Leu Leu Gly Phe Gly Val Ala Leu Ala Ser Leu  
595 600 605

atc gag aay tgt ccc aaa gac aac aag gac tgc agc tcc tac gcc agc  
1872

Ile Glu Lys Cys Pro Lys Asp Asn Lys Asp Cys Ser Ser Tyr Gly Ser  
610 615 620

ttc agt agc gca gta gtg ctg gaa ctg ttc aag ctg acc ata ggc ctg ggt  
1920

Phe Ser Asp Ala Val Leu Glu Leu Phe Lys Leu Thr Ile Gly Leu Gly  
625 630 635 640

gat ctg aac atc cag cag aac tcc aag tat ccc att ctg ttt ctg ttc  
1968

Asp Leu Asn Ile Gln Gln Asn Ser Lys Tyr Pro Ile Leu Phe Leu Phe  
645 650 655

ctg ctg atc acc tat gtc atc ctg acc ttt gtt ctg ctg acc aac atg  
2016

Leu Leu Ile Thr Tyr Val Ile Leu Thr Phe Val Leu Leu Leu Asn Met  
660 665 670

ctc att gct ctg atg ggc gag act gtg gag aac gtc tcc aag gag agc  
2064

Leu Ile Ala Leu Met Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser  
675 680 685

gaa cgc atc tgg cgc ctg cag aga gcc agg acc atc ttg gag ttt gag  
2112

Glu Arg Ile Trp Arg Leu Gln Arg Ala Arg Thr Ile Leu Glu Phe Glu  
690 695 700

aaa atg tta cca gaa tgg ctg agg agc aga ttc cgg atg gga gag ctg  
2160

Lys Met Leu Pro Glu Trp Leu Arg Ser Arg Phe Arg Met Gly Glu Leu  
710 715 720

tgc aaa gtg gcc gag gat gat ttc cga ctg tgt ttg cgg atc aat gag  
2208

Cys Lys Val Ala Glu Asp Phe Arg Leu Cys Leu Arg Ile Asn Glu  
725 730 735

gtg aag tgg act gaa  
2223

Val Lys Trp Thr Glu  
740

<210> 4  
<211> 3042  
<212> DNA  
<213> Homo sapiens

<220>  
<221> COS  
<222> (1)...(3039)

<400> 4  
atg gtt gga gga tgc agg tgg aca gaa gac gtg gag cct gca gaa gta  
Met Val Gly Gly Cys Arg Trp Thr Glu Asp Val Glu Pro Ala Glu Val  
1 5 10 15

aag gaa aag atg tcc ttt cgg gca gcc agg ctg agc atg agg aac aga  
Lys Glu Lys Met Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg  
20 25 30 35 36

agg aat gac act ctg gac agc acc cgg acc ctg tac tcc agc ggc tct 144  
 Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Tyr Ser Ser Ala Ser 35  
 40  
 45  
 cgg agc aca gac ttg tct tac agt gaa agc gac ttg gty aat ttt att 192  
 Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile 50  
 55  
 60  
 caa gca aat ttt aag aaa cga gaa tgt gtc ttc ttt acc aaa gat tcc 240  
 Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser 65  
 70  
 75  
 80  
 aag gcc acg gag aat gty tgc aag tgt ggc tat gcc cag agc cag cac 288  
 Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His 85  
 90  
 95  
 atg gaa gcc acc cag atc aac caa agt gag aaa tgg aac tac aag aaa 336  
 Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys 100  
 105  
 110  
 cac acc aag gaa ttt cct acc gac gcc ttt ggg gat att cag ttt gag 384  
 His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu 115  
 120  
 125  
 aca ctg ggg aag aaa ggg aag tat ata cgt ctg tcc tgc gac acg gac 432  
 Thr Leu Gly Lys Lys Gly Tyr Ile Arg Leu Ser Cys Asp Thr Asp 130  
 135  
 140  
 gcg gaa atc ctt tac gag ctg ctg acc cag cac tgg cac ctg aaa aca 480  
 Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr 145  
 150  
 155  
 ccc aac ctg gtc att tct gty acc ggg ggc gcc aag aac ttc gcc ctg 528  
 Pro Asn Leu Val Ile Ser Val Thr Gly Ala Lys Asn Phe Ala Leu 160  
 165  
 170  
 aag cgg cgc atg cgc aag atc ttc agc cgg ctg atc ttc atc ggc cag 576  
 Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln 180  
 185  
 190  
 tcc aaa ggt gct tgg att ctg acg gga gcc acc cat tat ggc ctg atg 624  
 Ser Lys Gly Ala Trp Ile Leu Thr Gly Tyr His Tyr Gly Leu Met 195  
 200  
 205  
 aag tac ctg ggg gag gty gty aga gat aac acc atc agc agg agt tca 672  
 Lys Tyr Leu Glu Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser 210  
 215  
 220  
 gag gag aat att gty gcc att ggc ala gca gct tgg ggc atg gtc tcc 720  
 Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser 225  
 230  
 235  
 240  
 aac cgg gac acc ctg atc agc aat tgc gat gct gag ggc tat ttt tta 768  
 Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu 245  
 250  
 255  
 gcc cag tac ctt atg gat gac ttc aca aga gat cca ctg tat atc ctg 816  
 Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Tyr Ile Leu 260  
 265  
 270

gac aac aac cac aca cat ttg ctg ctg gac aat ggc tgt cat gga 864  
 Asp Asn Asn His Thr His Leu Leu Val Asp Asn Gly Cys His Gly 275  
 280  
 285  
 cat ccc act gtc gaa gca aag ctg cag aat cag cta gag aag tat atc 912  
 His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile 290  
 295  
 300  
 tct gag cgc act att caa gat tcc aac tat ggt ggc aag atc ccc att 960  
 Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Ile Pro Ile 305  
 310  
 315  
 320  
 gty tgt ttt gcc caa gga ggt gga aaa gag act ttg aaa gcc atc aat  
 1008  
 Val Cys Phe Ala Gln Gly Gly Lys Lys Glu Thr Leu Lys Ala Ile Asn 325  
 330  
 335  
 acc tcc atc aaa aat aaa att cct tgt gty gty gaa ggc tgc ggc 1056  
 Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Glu Gly Ser Gly 340  
 345  
 350  
 cag atc gct gat gty atc gct agc ctg gty gag gty gag gat gcc ctg  
 1104  
 Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Asp Ala Leu 355  
 360  
 365  
 aca tct tct gcc gtc aag gag aag ctg gty cgc ttt tta ccc cgc acg  
 1152  
 Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr 370  
 375  
 380  
 gty tcc cgg ctg cct gag gag gag act gag agt tgg atc aaa tgg ctg  
 1200  
 Val Ser Arg Leu Pro Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu 385  
 390  
 395  
 aaa gaa att ctg gaa tgt tct cac cta tta aca gtt att aaa atg gaa  
 1248  
 Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu 405  
 410  
 415  
 gaa gct ggg gat gaa att gty agc aat gcc atc tcc tac gct cta tac  
 1296  
 Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr 420  
 425  
 430  
 aaa gcc ttc agc acc agt gag caa gac aag gat aac tgg aat ggg cag  
 1344  
 Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln 435  
 440  
 445  
 ctg aag ctt ctg ctg gag tgg aac cag ctg gac tta gcc aat gat gag  
 1392  
 Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu 450  
 455  
 460  
 att ttc acc aat gac cgc cga tgg gag aag agc aaa ccg agc ctg aga  
 1440  
 Ile Phe Thr Asn Asp Arg Arg Trp Glu Lys Ser Lys Pro Arg Leu Arg 465  
 470  
 480

gag aca ata atc cag gtc aca tgg ctg gaa aat ggt aga atc aag gtt  
 1488  
 Asp Thr Ile Ile Gln Val Thr Trp Leu Glu Asn Gly Arg Ile Lys Val  
 485 490 495

gag agc aaa gat gtg act gac ggc aaa gcc tct tct cat atg ctg gtg  
 1536  
 Glu Ser Lys Asp Val Thr Asp Gly Lys Ala Ser Ser His Met Leu Val  
 500 505 510

gtt ctc aag tct gct gac ctt caa gaa gtc atg ttt acg gct ctc ata  
 1584  
 Val Leu Lys Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile  
 515 520 525

aag gac aga ccc aag ttt gtc cgc ctc ttt ctg gag aat ggc ttg aac  
 1632  
 Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn  
 530 535 540

cta cgg aag ttt ctc acc cat gat gtc ctc act gaa ctc ttc tcc aac  
 1680  
 Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn  
 545 550 555 560

cac ttc agc acg ctt gtg tac cgg aat ctg cag atc gcc aag aat tcc  
 1728  
 His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser  
 565 570 575

tat aat gat gcc ctc ctc acg ttt gtc tgg aaa ctg gtt ggc aac ttc  
 1776  
 Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe  
 580 585 590

cga aga ggc ttc cgg aag gaa gac aga aat ggc cgg gac gag atg gac  
 1824  
 Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp  
 595 600 605

ata gaa ctc cac gac gtg tct cct att act cgg cac ccc ctg caa gct  
 1872  
 Ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala  
 610 615 620

ctc ttc atc tgg gcc att ctt cag aat aag aag gaa ctc tcc aaa gtc  
 1920  
 Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val  
 625 630 635 640

att tgg gag cag acc agg ggc tgc act ctg gca gcc ctg gga gcc agc  
 1968  
 Ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser  
 645 650 655

aag ctt ctg aag act ctg gcc aaa gtg aag aac gac atc aat gct gct  
 2016  
 Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala  
 660 665 670

ggg gag tcc gag gag ctg gct aat gag tac ctg acc cgg gct gtt ggt  
 2064  
 Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Leu Thr Arg Ala Val Gly  
 675 680 685

gag tcc aca gtg tgg aat gct gtg ggc ggc gat ctg cca tgt ggc  
 2112  
 Glu Ser Thr Val Trp Asn Ala Val Val Gly Ala Asp Leu Pro Cys Gly  
 690 695 700

aca gac att gcc agc ggc act cat aga cca gat ggt gga gag ctg ttc  
 2160  
 Thr Asp Ile Ala Ser Gly Thr His Arg Pro Asp Gly Gly Glu Leu Phe  
 705 710 715 720

act gag tgt tac agc agc gat gaa gac ttg gca gaa cag ctg ctg gtc  
 2208  
 Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val  
 725 730 735

tat tcc tot gaa gct tgg ggt gga agc aac tgt ctg gag ctg ggc gtg  
 2256  
 Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val  
 740 745 750

gag gcc aca gac cag cat ttc atc gcc cag cct ggg gtc cag aat ttt  
 2304  
 Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe  
 755 760 765

ctt tct aag caa tgg tat gga gag att tcc cga gac acc aag aac tgg  
 2352  
 Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp  
 770 775 780

aag att atc ctg tgt ctg ttt att ata ccc ttg gtg ggc tgt ggc ttt  
 2400  
 Lys Ile Ile Leu Cys Leu Phe Phe Ile Ile Pro Leu Val Gly Cys Gly Phe  
 785 790 795 800

gta tca ttt agg aag aaa cct gtc gac aag cac aag aag ctg ctt tgg  
 2448  
 Val Ser Phe Arg Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp  
 805 810 815

tac tat gtg gcg ttc ttc acc tcc ccc ttc gtg gtc ttc tcc tgg aat  
 2496  
 Tyr Tyr Val Ala Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn  
 820 825 830

gtg gtc ttc tac atc gcc ttc ctc ctg ctg ttt gcc tac gtg ctg ctc  
 2544  
 Val Val Phe Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu  
 835 840 845

atg gat ttc cat tgg gtg cca cac ccc ccc ccc gag ctg gtc ctg tac tgg  
 2592  
 Met Asp Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser  
 850 855 860

16

ctg gtc ttt gtc ctc ttc tgt gat gaa gtg aga cag ggc cgg cgt  
 2640  
 Leu Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Gly Arg Pro Ala  
 865 870 875 880

gct ccc agt gcg ggg ccc gcc aag ccc acg ccc acc cgg aac tcc atc  
 2688  
 Ala Pro Ser Ala Gly Pro Ala Lys Pro Thr Thr Arg Asn Ser Ile  
 885 890 895

tgg ccc gca agc tcc aca cgc agc ccc ggt tcc cgc tca cgc cac tcc  
 2736  
 Trp Pro Ala Ser Ser Thr Arg Ser Pro Gly Ser Arg Ser Arg His Ser  
 900 905 910

ttc cac act tcc ctg caa gct gag ggt gcc agc tct ggc ctt ggc cag  
 2784  
 Phe His Thr Ser Leu Gln Ala Glu Gly Ala Ser Ser Gly Leu Gly Gln  
 915 920 925

ccc aga aag ggg ctc cca cag tgc agc ggt ggg ctg aag ggc tcc tca  
 2832  
 Pro Arg Lys Gly Leu Pro Gln Cys Ser Gly Gly Leu Lys Gly Ser Ser  
 930 935 940

agt gcc gcc aaa gtg gga gcc cag gca gag gag gty ccg aga gca agc  
 2880  
 Ser Ala Ala Lys Val Gly Ala Gln Ala Glu Val Pro Arg Ala Ser  
 945 950 955 960

ggg ggc tgt gag gac tgc cag cac gct gtc acc tct cag aag cgt aag  
 2928  
 Glu Gly Cys Glu Asp Cys Gln His Ala Val Thr Ser Gln Lys Arg Lys  
 965 970 975

aca gca atg gac caa aca gac gaa gat ctc ttc ccc tat gga gca ttc  
 2976  
 Thr Ala Met Asp Gln Thr Asp Glu Asp Leu Phe Pro Tyr Gly Ala Phe  
 980 985 990

tac cag ttc ctg atg att tcc aag agc ttc cga gga gag gag atg agc  
 3024  
 Tyr Gln Phe Leu Met Ile Ser Arg Ser Phe Arg Gly Glu Glu Met Ser  
 995 1000 1005

atc ggc aag cag cac taa  
 3042  
 Ile Gly Lys Gln His  
 1010

<210> 5  
 <211> 1013  
 <212> PRT  
 <213> Homo sapiens

<400> 5  
 Met Val Gly Gly Cys Arg Trp Thr Glu Asp Val Glu Pro Ala Glu Val  
 1 5 10 15  
 Lys Glu Lys Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg  
 20 25 30

17

Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser  
 35 40 45  
 Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile  
 50 55 60  
 Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp Ser  
 65 70 75 80  
 Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser Gln His  
 85 90 95  
 Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn Tyr Lys Lys  
 100 105 110  
 His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp Ile Gln Phe Glu  
 115 120 125  
 Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser Cys Asp Thr Asp  
 130 135 140  
 Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp His Leu Lys Thr  
 145 150 155 160  
 Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys Asn Phe Ala Leu  
 165 170 175  
 Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln  
 180 185 190  
 Ser Lys Gly Ala Trp Ile Leu Thr Gly Thr His Tyr Gly Leu Met  
 195 200 205  
 Lys Tyr Leu Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser  
 210 215 220  
 Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser  
 225 230 235 240  
 Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu  
 245 250 255  
 Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Tyr Ile Leu  
 260 265 270  
 Asp Asn Asn His Thr His Leu Leu Leu Val Asp Asn Gly Cys His Gly  
 275 280 285  
 His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile  
 290 295 300  
 Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Lys Ile Pro Ile  
 305 310 315 320  
 Val Cys Phe Ala Gln Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn  
 325 330 335  
 Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Glu Gly Ser Gly  
 340 345 350  
 Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu  
 355 360 365  
 Thr Ser Ser Ala Val Lys Glu Lys Leu Val Arg Phe Leu Pro Arg Thr  
 370 375 380  
 Val Ser Arg Leu Pro Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu  
 385 390 395 400  
 Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu  
 405 410 415  
 Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr  
 420 425 430  
 Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln  
 435 440 445  
 Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu  
 450 455 460  
 Ile Phe Thr Asn Asp Arg Arg Trp Glu Lys Ser Lys Pro Arg Leu Arg  
 465 470 475 480  
 Asp Thr Ile Ile Gln Val Thr Trp Leu Glu Asn Gly Arg Ile Lys Val  
 485 490 495  
 Glu Ser Lys Asp Val Thr Asp Gly Lys Ala Ser Ser His Met Leu Val  
 500 505 510

Val	Leu	Lys	Ser	Ala	Asp	Leu	Gln	Glu	Val	Met	Phe	Thr	Ala	Leu	Ile
515	515	515	515	515	515	515	515	515	515	515	515	515	515	515	515
Lys	Asp	Arg	Pro	Lys	Phe	Val	Arg	Leu	Phe	Leu	Glu	Asn	Gly	Leu	Asn
530	530	530	530	530	530	530	530	530	530	530	530	530	530	530	530
Leu	Arg	Lys	Phe	Leu	Thr	His	Asp	Val	Leu	Thr	Glu	Leu	Phe	Ser	Asn
545	545	545	545	545	545	545	545	545	545	545	545	545	545	545	545
His	Phe	Ser	Thr	Leu	Val	Tyr	Arg	Asn	Leu	Gln	Ile	Ala	Lys	Asn	Ser
565	565	565	565	565	565	565	565	565	565	565	565	565	565	565	565
Tyr	Asn	Asp	Ala	Leu	Leu	Thr	Phe	Val	Trp	Lys	Leu	Val	Ala	Asn	Phe
580	580	580	580	580	580	580	580	580	580	580	580	580	580	580	580
Arg	Arg	Gly	Phe	Arg	Lys	Glu	Asp	Arg	Asn	Gly	Arg	Asp	Glu	Met	Asp
595	595	595	595	595	595	595	595	595	595	595	595	595	595	595	595
Ile	Glu	Leu	His	Asp	Val	Ser	Pro	Ile	Thr	Arg	His	Pro	Leu	Gln	Ala
610	610	610	610	610	610	610	610	610	610	610	610	610	610	610	610
Leu	Phe	Ile	Trp	Ala	Ile	Leu	Gln	Asn	Lys	Lys	Glu	Leu	Ser	Lys	Val
625	625	625	625	625	625	625	625	625	625	625	625	625	625	625	625
Ile	Trp	Glu	Gln	Thr	Arg	Gly	Cys	Thr	Leu	Ala	Ala	Leu	Gly	Ala	Ser
645	645	645	645	645	645	645	645	645	645	645	645	645	645	645	645
Lys	Leu	Leu	Lys	Thr	Leu	Ala	Lys	Val	Lys	Asn	Asp	Ile	Asn	Ala	Ala
660	660	660	660	660	660	660	660	660	660	660	660	660	660	660	660
Gly	Glu	Ser	Glu	Glu	Leu	Ala	Asn	Glu	Tyr	Leu	Thr	Arg	Ala	Val	Gly
675	675	675	675	675	675	675	675	675	675	675	675	675	675	675	675
Glu	Ser	Thr	Val	Trp	Asn	Ala	Val	Val	Gly	Ala	Asp	Leu	Pro	Cys	Gly
690	690	690	690	690	690	690	690	690	690	690	690	690	690	690	690
Thr	Asp	Ile	Ala	Ser	Gly	Thr	His	Arg	Pro	Asp	Gly	Gly	Glu	Leu	Phe
705	705	705	705	705	705	705	705	705	705	705	705	705	705	705	705
Thr	Glu	Cys	Tyr	Ser	Ser	Asp	Glu	Asp	Leu	Ala	Glu	Gln	Leu	Leu	Val
725	725	725	725	725	725	725	725	725	725	725	725	725	725	725	725
Tyr	Ser	Cys	Glu	Ala	Trp	Gly	Gly	Ser	Asn	Cys	Leu	Glu	Leu	Ala	Val
740	740	740	740	740	740	740	740	740	740	740	740	740	740	740	740
Glu	Ala	Thr	Asp	Gln	His	Phe	Ile	Ala	Gln	Pro	Gly	Val	Gln	Asn	Phe
755	755	755	755	755	755	755	755	755	755	755	755	755	755	755	755
Leu	Ser	Lys	Gln	Trp	Tyr	Gly	Glu	Ile	Ser	Arg	Asp	Thr	Lys	Asn	Trp
770	770	770	770	770	770	770	770	770	770	770	770	770	770	770	770
Lys	Ile	Ile	Leu	Cys	Leu	Phe	Ile								

Tyr	Gln	Phe	Leu	Met	Ile	Ser	Arg	Ser	Phe	Arg	Gly	Glu	Met	Ser
995								1000					1005	
Ile	Gly	Lys	Gln	His										
1010														
<210>	6													
<221>	3039													
<212>	DNA													
<213>	Homo sapiens													
<220>														
<221>	CDS													
<222>	(1)...(3039)													
<400>	6													
atg	gtt	gga	gga	tgc	aag	tgg	aca	gaa	gac	gtg	gag	cct	gca	gaa gta
1				5						10				15
aag	gaa	aag	atg	tcc	ttt	cgg	gca	gcc	agg	ctc	agc	atg	agg	aac aga
	Lys	Glu	Lys	Met	Ser	Phe	Arg	Ala	Ala	Arg	Leu	Ser	Met	Arg Asn Arg
			20			25					30			
agg	aat	gac	act	ctg	gac	agc	acc	cgg	acc	ctg	tac	tcc	agc	gcg tct
	Arg	Ans	Asp	Thr	Leu	Asp	Ser	Thr	Arg	Thr	Leu	Tyr	Ser	Ser Ala Ser
	35				40						45			
cgg	agc	aca	gac	ttg	tct	tac	agt	gaa	agc	gac	ttg	gtg	aat	ttt att
	Arg	Ser	Thr	Asp	Leu	Ser	Tyr	Ser	Glu	Ser	Asp	Leu	Val	Ans Phe Ile
	50					55					60			
caa	gca	aat	ttt	aag	aaa	cga	gaa	tgt	gtc	ttc	ttt	acc	aaa	gat tcc
	Gln	Ala	Asn	Phe	Lys	Lys	Arg	Glu	Cys	Val	Phe	Phe	Thr	Lys Asp Ser
	65				70					75				80
aag	gcc	acg	gag	aat	gtg	tgc	aag	tgt	ggc	tat	gcc	cag	agc	cag cac
	Lys	Ala	Thr	Glu	Asn	Val	Cys	Lys	Cys	Gly	Tyr	Ala	Gln	Ser Gln His
					85				90					95
atg	gaa	ggc	acc	cag	atc	aac	caa	agt	gag	aaa	tgg	aac	tac	aag aaa
	Met	Glu	Gly	Thr	Glu	Ile	Asn	Gln	Ser	Glu	Lys	Trp	Asn	Tyr Lys Lys
			100					105			110			
cac	acc	aag	gaa	ttt	cct	acc	gac	gcc	ttt	ggg	gat	att	cag	ttt gag
	His	Thr	Lys	Glu	Phe	Pro	Thr	Asp	Ala	Phe	Gly	Asp	Ile	Gln Phe Glu
			115			120					125			
aca	ctg	ggg	aag	aaa	ggg	aag	tat	ata	cgt	ctg	tcc	tgc	gac	acg gac
	Thr	Leu	Gly	Lys	Lys	Gly	Lys	Tyr	Ile	Arg	Leu	Ser	Cys	Asp Thr Asp
			130				135				140			
gcg	gaa	atc	ctt	tac	gag	ctg	ctg	acc	cag	cac	tgg	cac	ctg	aaa aca
	Ala	Glu	Ile	Leu	Tyr	Glu	Leu	Leu	Thr	Gln	His	Trp	His	Leu Lys Thr
			145			150				155				160
ccc	aac	ctg	gtc	att	tct	gtg	acc	ggg	ggc	gcc	aag	aac	ttc	gcc ctg
	Pro	Ans	Leu	Val	Ile	Ser	Val	Thr	Gly	Gly	Ala	Lys	Asn	Phe Ala Leu
					165				170					175

aag ccg cgc atg cgc aag atc ttc agc cgg ctc atc tac atc ggc cag 576  
 Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile Tyr Ile Ala Gln  
 180 185 190  
 tcc aaa ggt gct tgg atc ctc acg gga ggc acc cat tat ggc ctg atg 624  
 Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His Tyr Gly Leu Met  
 195 200 205  
 aag tac ctc ggg gag ctg aga gat aac acc atc agc agg agt tca 672  
 Lys Tyr Leu Gly Glu Val Val Arg Asp Asn Thr Ile Ser Arg Ser Ser  
 210 215 220  
 gag gag aat att ctg gcc att ggc ata gca gct tgg ggc atg gtc tcc 720  
 Glu Glu Asn Ile Val Ala Ile Gly Ile Ala Trp Gly Met Val Ser  
 225 230 235 240  
 aac cgg gac acc ctc atc aag aat tgc gat gct gag ggc tat ttt tta 768  
 Asn Arg Asp Thr Leu Ile Arg Asn Cys Asp Ala Glu Gly Tyr Phe Leu  
 245 250 255  
 gcc cag tac ctt atg gat gac ttc aca aga gat cca ctg tat atc ctg 816  
 Ala Gln Tyr Leu Met Asp Asp Phe Thr Arg Asp Pro Leu Tyr Ile Leu  
 260 265 270  
 gac aac aac cac aca cat ttg ctg ctc ctg gac aat ggc tgt cat gga 864  
 Asp Asn Asn His Thr His Leu Leu Val Val Asp Asn Gly Cys His Gly  
 275 280 285  
 cat ccc act gtc gaa gca aag ctc cgg aat cag cta gag aag tat atc 912  
 His Pro Thr Val Glu Ala Lys Leu Arg Asn Gln Leu Glu Lys Tyr Ile  
 290 295 300  
 tct gag cgc act att caa gat tcc aac tat ggt ggc aag atc ccc att 960  
 Ser Glu Arg Thr Ile Gln Asp Ser Asn Tyr Gly Gly Ile Pro Ile  
 305 310 315 320  
 gtg tgt ttt gcc caa gga ggt gga aaa gag act ttg aaa gcc atc aat  
 1008  
 Val Cys Phe Ala Gln Gly Gly Lys Glu Thr Leu Lys Ala Ile Asn  
 325 330 335  
 acc tcc atc aaa aat aaa att cct tgt gtg gtg gaa ggc tgg ggc  
 1056  
 Thr Ser Ile Lys Asn Lys Ile Pro Cys Val Val Val Glu Gly Ser Gly  
 340 345 350  
 cag atc gct gat gtg atc gct agc ctg gtg gag gtg gag gat ggc ctg  
 1104  
 Gln Ile Ala Asp Val Ile Ala Ser Leu Val Glu Val Glu Asp Ala Leu  
 355 360 365  
 aca tct tct gcc gtc aag gag aag ctg gtg cgc ttt tta ccc cgc acg  
 1152  
 Thr Ser Ser Ala Val Lys Lys Leu Val Arg Phe Leu Pro Arg Thr  
 370 375 380  
 gtg tcc cgg ctg cct gag gag gag act gag agt tgg atc aaa tgg ctc  
 1200  
 Val Ser Arg Leu Leu Glu Glu Glu Thr Glu Ser Trp Ile Lys Trp Leu  
 385 390 395 400

aag gaa att ctc gaa tgt tct cac cta tta aca gtt att aaa atg gaa  
 1248  
 Lys Glu Ile Leu Glu Cys Ser His Leu Leu Thr Val Ile Lys Met Glu  
 405 410 415  
 gaa gct ggg gat gaa att gtg agc aat gcc atc tcc tac gct cta tac  
 1296  
 Glu Ala Gly Asp Glu Ile Val Ser Asn Ala Ile Ser Tyr Ala Leu Tyr  
 420 425 430  
 aac gcc ttc agc acc agt gag caa gac aag gat aac tgg aat ggg cag  
 1344  
 Lys Ala Phe Ser Thr Ser Glu Gln Asp Lys Asp Asn Trp Asn Gly Gln  
 435 440 445  
 ctg aag ctt ctg ctg gag tgg aac cag ctg gac tta gcc aat gat gag  
 1392  
 Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu Ala Asn Asp Glu  
 450 455 460  
 att ttc acc aat gac cgc cga tgg gag aag agc aaa ccg agg ctc aga  
 1440  
 Ile Phe Thr Asn Asp Arg Arg Trp Glu Lys Ser Lys Pro Arg Leu Arg  
 465 470 475 480  
 gac aca ata atc cag gtc aca tgg ctg gaa aat ggt aga atc aag gtt  
 1488  
 Asp Thr Ile Ile Gln Val Thr Trp Leu Glu Asn Gly Arg Ile Lys Val  
 485 490 495  
 gag agc aag gat gtg act gac ggc aac gcc tct tct cat atg ctg gtg  
 1536  
 Glu Ser Lys Asp Val Thr Asp Gly Lys Ala Ser Ser His Met Leu Val  
 500 505 510  
 gtt ctc aag tct gct gac ctt caa gaa gtc atg ttt acy gct ctc ata  
 1584  
 Val Leu Lys Ser Ala Asp Leu Gln Glu Val Met Phe Thr Ala Leu Ile  
 515 520 525  
 aag gac aga ccc aag ttt gtc cgc ctc ttt ctg gag aat ggc ttg aac  
 1632  
 Lys Asp Arg Pro Lys Phe Val Arg Leu Phe Leu Glu Asn Gly Leu Asn  
 530 535 540  
 cta cgg aag ttt ctc acc cat gat gtc ctc act gaa ctc ttc tcc aac  
 1680  
 Leu Arg Lys Phe Leu Thr His Asp Val Leu Thr Glu Leu Phe Ser Asn  
 545 550 555 560  
 cac ttc agc acg ctt gtg tac cgg aat ctg cag atc gcc aag aat tcc  
 1728  
 His Phe Ser Thr Leu Val Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser  
 565 570 575  
 tat aat gat gcc ctc ctc acg ttt gtc tgg aaa ctg gtt gcg aac ttc  
 1776  
 Tyr Asn Asp Ala Leu Leu Thr Phe Val Trp Lys Leu Val Ala Asn Phe  
 580 585 590 595

cga aga ggc ttc cgg aag gaa gac aga aat ggc cgg gac gag atg gac  
1824  
Arg Arg Gly Phe Arg Lys Glu Asp Arg Asn Gly Arg Asp Glu Met Asp  
595 600  
ata gaa ctc cac gac gtg tct cct att act cgg cac ccc ctg caa gct  
1872  
ile Glu Leu His Asp Val Ser Pro Ile Thr Arg His Pro Leu Gln Ala  
610 615 620  
ctc ttc atc tgg gcc att ctt cag aat aag aag gaa ctc tcc aaa gtc  
1920  
Leu Phe Ile Trp Ala Ile Leu Gln Asn Lys Lys Glu Leu Ser Lys Val  
625 630 635 640  
att tgg gag cag acc agg ggc tgc act ctg gca gcc ctg gga gcc agc  
1968  
ile Trp Glu Gln Thr Arg Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser  
645 650 655  
aag ctt ctg aag act ctg gcc aaa gtg aag aac gac atc aat gct gct  
2016  
Lys Leu Leu Lys Thr Leu Ala Lys Val Lys Asn Asp Ile Asn Ala Ala  
660 665 670  
ggg gag tcc gag gag ctg gct aat gag tac ctg acc cgg gct gtt ggt  
2064  
Gly Glu Ser Glu Glu Leu Ala Asn Glu Tyr Leu Thr Arg Ala Val Gly  
675 680 685  
gag tcc aca gtg tgg aat gct gtg ggc ggc gat ctg cca tgt ggc  
2112  
Glu Ser Thr Val Trp Asn Ala Val Val Gly Ala Asp Leu Pro Cys Gly  
690 695 700  
aag gac att gcc agc ggc act cat aga cca gat ggt gga gag ctg ttc  
2160  
Thr Asp Ile Ala Ser Gly Thr His Arg Pro Asp Gly Gly Glu Leu Phe  
705 710 715 720  
act gag tgt tac agc agc gat gaa gac ttg gca gaa cag ctg ctg gtc  
2208  
Thr Glu Cys Tyr Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val  
725 730 735  
tat tcc tgt gaa gct tgg ggt gga agc aac tgt ctg gag ctg gcg gtg  
2256  
Tyr Ser Cys Glu Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val  
740 745 750  
gag gcc aca gac cag cat ttc atc gcc cag cct ggg gtc cag aat ttt  
2304  
Glu Ala Thr Asp Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe  
755 760 765  
ctt tct aag caa tgg tat gga gag att tcc cga gac acc aag aac tgg  
2352  
Leu Ser Lys Gln Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp  
770 775 780

aag att atc ctg tgt ctg ttt att ata ccc ttg gtg ggc tgt ggc ttt  
2400  
Lys Ile Ile Leu Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe  
785 790 795 800  
gta tca ttt agg aag aaa cct gtc gac aag cac aag aag ctg ctt tgg  
2448  
Val Ser Phe Arg Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp  
805 810 815  
tac tat gtg gcg ttc ttc acc tcc ccc ttc gtg gtc ttc tcc tgg aat  
2496  
Tyr Tyr Val Ala Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn  
820 825 830  
gtg gtc ttc tac atc gcc ttc ctc ctg ctg ttt gcc tac gtg ctg ctc  
2544  
Val Val Phe Tyr Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu  
835 840 845  
atg gat ttc cat tcg gtg cca cac ccc ccc gag ctg gtc ctg tac tcc  
2592  
Met Asp Phe His Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser  
850 855 860  
ctg gtc ttt gtc ctc ttc tgt gat gaa gtg aga cag ggc cgg cgt gct  
2640  
Leu Val Phe Val Leu Phe Cys Asp Glu Val Arg Gln Gly Arg Pro Ala  
865 870 875 880  
gct ccc agt gcg ggg ccc gcc ccc agc ccc acc cgg aac tcc atc  
2688  
Ala Pro Ser Ala Gly Pro Ala Lys Pro Thr Pro Thr Arg Asn Ser Ile  
885 890 895  
tgg ccc gca agc tcc aca cgc agc ccc ggt tcc cgc tca cgc cac tcc  
2736  
Trp Pro Ala Ser Ser Thr Arg Ser Pro Gly Ser Arg Ser Arg His Ser  
900 905 910  
ttc cac act tcc ctg caa gct gag ggt gcc agc tct ggc ctt ggc cag  
2784  
Phe His Thr Ser Leu Gln Ala Glu Gly Ala Ser Ser Gly Leu Gly Gln  
915 920 925  
ccc aga aag ggg ctc cca cag tgc agc ggt ggg ctg aag ggc tcc tca  
2832  
Pro Arg Lys Gly Leu Pro Gln Cys Ser Gly Gly Leu Lys Gly Ser Ser  
930 935 940  
agt gcc gcc aaa gtg gga gcc cag gca gag gag gtg ccg aga gca agc  
2880  
Ser Ala Ala Lys Val Gly Ala Gln Ala Glu Glu Val Pro Arg Ala Ser  
945 950 955 960  
gag ggc tgt gag gac tgc cag cac gct gtc acc tct cag aag cgt aag  
2928  
Glu Gly Cys Glu Asp Cys Gln His Ala Val Thr Ser Gln Lys Arg Lys  
965 970 975



aca gca atg gac caa aca gac gaa gat ctc ttc ccc tat gga gca ttc  
2976  
Thr Ala Met Asp Gln Thr Asp Gln Asp Leu Phe Pro Tyr Gly Ala Phe  
980  
tac cag ttc ctg atg att tcc agg agc ttg cga gga gag gag atg agc  
3024  
Tyr Gln Phe Leu Met Ile Ser Arg Ser Phe Arg Gly Gln Met Ser  
995  
atc ggc aag cag cac  
3039  
Ile Gly Lys Gln His  
1010

**THIS PAGE BLANK (USPTO)**

This Page is inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox**

**THIS PAGE BLANK (USPTO)**